UNIVERSITY OF CAPE TOWN

DEPARTMENT OF MATHEMATICAL STATISTICS

****

A TIME SERIES APPROACH TO THE MONETARY SECTOR OF THE SOUTH AFRICAN ECONOMY

by

C.H. Dietzsch

****

A thesis prepared under the supervision of Professor C.G. Troskie of the Department of Mathematical Statistics in partial fulfilment of the requirements for the degree of Master of Science in Mathematical Statistics

To my parents

P R E F A C E

This thesis provides an investigation of the applicability of time series analysis to the process of economic model building. Chapter 1 explains the position of the Box-Jenkins approach to time series analysis in relation to other techniques of analysis. In Chapters 2 and 3 the theory of model building is discussed. In Chapter 4 an econometric model is analysed in detail from a time series approach.

I would like to thank Professor C.G. Troskie for his supervision. In addition I would like to thank Dr. A.M. Hurwitz for his encouragement and assistance and making available to me previous research done. Many thanks to Mrs. Cousins for her superb typing and Mr. G.D.I. Barr for his criticisms during the final writing up of this thesis.

I am indebted to the Council for Scientific and Industrial Research for the scholarship awarded which enabled me to undertake the research necessary for the completion of this thesis.

C O N T E N T S

CHAPTER 1

THE METHODOLOGY OF ECONOMIC TIME SERIES MODELLING

1.1 INTRODUCTION

Economic data are analysed to obtain an understanding of the behaviour of an economic system and to produce forecasts for certain variables. Two different approaches can be followed to reach the above goals. The economic approach lends itself more to the description of the behaviour of an economic system. Relationships between the variables are established and are finally grouped into a model. These relationships can be solved individually or simultaneously. In econometric analysis the behaviour of a variable is usually only described by the influence of other explanatory variables. The theory of time series analysis is built around the characteristics of the variable itself, for example the presence of a trend or a seasonal effect. The purpose of the analysis of a single time series is to produce forecasts which could be used for control or planning purposes. Forecasts are produced on the assumption that the structure of an economic system will not change in the near future.

The analysis of time series has been extended to the bivariate case where the influence of an explanatory variable is

taken into account. The approach can be extended to the multivariate case so that one has the complement of a simultaneous econometric model. A lack of the necessary programming facilities has limited practical applications.

As econometric modelling and time series modelling are based on different concepts, it is debatable to which method of analysis one should give preference. One should be cautious to compare results of these different philosophies and there can be more benefit if research done from these fields is used complementary.

At the University of Cape Town a considerable amount of research has been done on econometric modelling. The author decided to follow the time series approach to economic modelling with the aim to employ useful results in future economic analysis together with those obtained with econometric modelling.

1.2 APPROACHES IN TIME SERIES ANALYSIS

A time series can be analysed in the frequency or the time domain. An analysis in the frequency domain is based on the power spectrum and is generally known as spectral analysis. A curve of the power spectrum shows how the variance of a stochastic process is distributed with frequency. In the time domain the basic tool for analysis is the autocorrelation function which is interrelated with the power spectrum. The

power spectrum is the Fourier transform of the autocorrelation function. Each function explains different characteristics of the time series so that these functions are complementary. Although usable characteristics of a time series is identified through spectral analysis, the author's primary interest is in the time domain field which links up more directly with econometric analysis. A large amount of techniques are available for analysis in the time domain such as exponential smoothing, adaptive filtering, Bayesian forecasting and ARIMA-modelling. The ARIMA models, also known as Box-Jenkins models, were of more interest as it allows the analyser to choose an appropiate model via specified strategies from a wide class of models. Thus it is possible that the model which actually generates a time series can be identified. Some of the above mentioned techniques can also be expressed as an ARIMA model. A further benefit is that the analyser knows the data and its characteristics better using the Box-Jenkins method as he has to exercise judgement at various stages in the model building procedure. The modelling procedure is based on sound distribution theory and can be extended to the multivariate case. The method also has disadvantages. It is not fully automatic as the analyser has to intervene during the model building process, while other methods only require the input of the original data and the results are produced automatically by the necessary programs. It is possible that the techniques can be incorrectly applied by the inexperienced user. A long series of data is always required which could eliminate the use of the Box-Jenkins method in certain practical cases when sufficient

data is not available.

1.3 TIME SERIES TERMINOLOGY

In this section we explain the different concepts to be used in time series analysis.

A time series implies a quantitative variable X_t that is ordered in time. X_t is a type of stochastic process as it is determined by probabilistic laws. A time series may be either continuous or discrete. If data can be recorded at every moment of time, then the data forms a continuous time series. A series is discrete if the data is given at specific instants of time. In nearly all economic time series the observations are recorded at equidistant points in time so that we only consider the discrete case.

The model building theory is built around the concept of stationarity as certain characteristics of a series X_t can be derived if stationarity is assumed. A time series is strictly stationary if the set of random variables $\{X_{t_1},\ldots,X_{t_p}\}$ has the same joint distribution as the set $\{X_{t_1+k},\ldots,X_{t_p+k}\}$ for all $t_1,\ldots,t_p$ and k. Thus

$$P[X_{t_1} \leq x_1,\ X_{t_2} \leq x_2,\ldots,X_{t_p} \leq x_p]$$
$$= P[X_{t_1+k} \leq x_1,\ X_{t_2+k} \leq x_2,\ldots,X_{t_p+k} \leq x_p] \qquad (1.1)$$

where P means the probability. The joint distribution of

$X_{t_1},\ldots,X_{t_p}$ depends only on the intervals between the time points $t_1,\ldots,t_p$ and not where the points are relative to the origin of the time axis.

As a very large amount of data is required to test a time series for strict stationarity, we define weak stationarity which will be adequate for practical purposes. For a time series to be weakly stationary, the $E[X_t]$ should be constant and $Cov[X_t, X_{t-r}] = \lambda_r$ for all r independent of time. λ_r is only a function of the time difference between X_t and X_{t-r}.

X_t is said to be a Gaussian process if the joint distribution of $X_{t_1},\ldots,X_{t_p}$ is multivariate normally distributed. For a Gaussian process the mean vector and the covariance matrix will be sufficient for a complete characterization of the distributional properties of the process. An alternative requirement for strict stationarity is that the mean and covariance of a Gaussian process must be independent of time.

The properties of a weakly stationary time series X_t can be usefully described by the autocovariance and autocorrelation functions. The autocovariance at lag k is defined as the covariance between X_t and X_{t-k}. Thus

$$\lambda_k = \text{Cov}[X_t, X_{t-k}]$$

$$= E[X_t X_{t-k}] - E[X_t]E[X_{t-k}]$$

If we assume that $E[X_t] = 0$ for all t, then

$$\lambda_k = E(X_t X_{t-k}] \tag{1.2}$$

Also $\lambda_{-k} = E[X_t X_{t+k}] = E[X_{t+k} X_t] = \lambda_k$

and $\lambda_0 = E[X_t^2] = \text{var}[X_t]$

The autocorrelation at lag k is defined as

$$\rho_k = \frac{E[X_t X_{t-k}]}{\sqrt{\text{var}(X_t)\text{var}(X_{t-k})}} = \frac{\lambda_k}{\lambda_0} \tag{1.3}$$

$$k = 0, \pm 1, \pm 2, \ldots$$

As the autocovariance and autocorrelation functions are symmetrical we only consider positive lags in future. A useful tool to be used together with the autocorrelation function is the partial autocorrelation function. The k^{th} partial autocorrelation is

$$\Phi_{kk} = \text{correlation } (X_t, X_{t-k} | X_{t-1}, \ldots, X_{t-k-1})$$

which is the correlation between X_t and X_{t-k} keeping the intermediate values fixed.

In a time series one only has one observation at each time t of the process. Assume X_t is a stationary process with $E[X_t] = \mu$.

Let $$M_{t_2-t_1+1}(X) = \frac{1}{t_2-t_1+1} \sum_{t=t_1}^{t_2} X_t$$

and $$M(X) = \lim_{\substack{t_1\to-\infty \\ t_2\to\infty}} M_{t_2-t_1+1}(X)$$

Birkhoff (1931) and Khintchin (1932) showed that

(i) $M(X)$ exists with probability one.

(ii) $M(X) = \mu$ if and only if $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \rho_k = 0$ and (i) holds.

The result justifies the mean of a process being estimated by taking the mean of successive values of a single realization. The process is called ergodic if $M(X) = \mu$ and $E[X_t-\mu]^2$ is finite. The above result of Birkhoff and Khintchin can be extended for an ergodic process implying that the autocorrelations can be estimated from a single observation at each time t.

Yule (1927) introduced the idea that a time series in which successive values are highly dependent can be regarded as generated from a series of independent random variables or "shocks" a_t. The shocks are random drawings from a fixed distribution with mean zero and constant variance σ_a^2. Thus for all t

$$E[a_t] = 0$$

and $var[a_t] = E[a_t^2] = \sigma_a^2$.

The autocovariance function is

$$\lambda_k = E[a_t\ a_{t-k}] = \sigma_a^2 \qquad \text{if} \quad k = 0$$
$$= 0 \qquad \text{if} \quad k > 0$$

The autocorrelation function is

$$\rho_k = 1 \qquad \text{if} \quad k = 0$$
$$= 0 \qquad \text{if} \quad k > 0$$

On the assumption that the fixed distribution is normally distributed, the sequence $a_t, a_{t-1}, \ldots$ is called a white noise process.

From the autocorrelation function certain behavioural patterns of the stochastic process can be identified. We require some plausible models with a recognisable pattern for the autocorrelation function. The white noise series has a very specific and easily recognisable shape. In mathematical form Yule's idea implies the following model:

$$X_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots \qquad (1.4)$$

for all t with μ the mean of the series X_t. This is known as a moving average model. The process X_t will be stationary if the sequence of θ's is finite or converges when infinite. From equation (1.4) we have the following equations

$$\tilde{X}_t = X_t - \mu = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \ldots$$
$$\tilde{X}_{t-1} = X_{t-1} - \mu = a_{t-1} - \theta_1 a_{t-2} - \theta_2 a_{t-3} - \ldots$$
$$\tilde{X}_{t-2} = X_{t-2} - \mu = a_{t-2} - \theta_1 a_{t-3} - \ldots$$
$$\vdots$$

where $\tilde{X}_t$ is the deviation of X_t from μ.

Applying specific weights $1, \phi_1, \phi_2, \ldots$ to each equation so that the coefficients of all positive lags in a_t are zero, then we obtain

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \phi_2 \tilde{X}_{t-2} + \ldots + a_t \qquad (1.5)$$

Thus $\tilde{X}_t$ is expressed as a linear function of previous values $\tilde{X}_{t-1}, \tilde{X}_{t-2}, \ldots$ and the random variable or error term a_t. This representation is known as an autoregressive model.

A further representation is obtained if we write $\tilde{X}_t$ as a linear function of previous values $\tilde{X}_{t-1}, \tilde{X}_{t-2}, \ldots$ and present and previous error terms $a_t, a_{t-1}, \ldots$. Thus

$$\tilde{X}_t = \phi_1 \tilde{X}_{t-1} + \phi_2 \tilde{X}_{t-2} + \ldots + a_t - \theta_1 a_{t-1} - \ldots \qquad (1.6)$$

which is known as the autoregressive-moving average representation.

A more compact mathematical representation can be obtained by introducing the backward shift operator B. B is defined so that

$$B\, X_t = X_{t-1}$$

and thus

$$B^k X_t = B^{k-1}BX_t = B^{k-1}X_{t-1} = X_{t-k}$$

The forward shift operator F is defined so that

$$F X_t = X_{t+1}$$

and thus

$$F^k X_t = X_{t+k}$$

F is the inverse of B.

Equation (1.4) can be written as

$$\begin{aligned} X_t &= \mu + a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \ldots \\ &= \mu + (1 - \theta_1 B - \theta_2 B^2 - \ldots)a_t \\ &= \mu + \theta(B) a_t \end{aligned} \qquad (1.7)$$

where

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots$$

Equation (1.5) becomes

$$\tilde{X}_t - \phi_1 B \tilde{X}_t - \phi_2 B \tilde{X}_t - \ldots = a_t$$

$$\therefore (1 - \phi_1 B - \phi_2 B^2 - \ldots)\tilde{X}_t = a_t$$

$$\therefore \phi(B)\tilde{X}_t = a_t \qquad (1.8)$$

where

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots$$

Using the above notation, equation (1.6) is

$$\phi(B)\tilde{X}_t = \theta(B)a_t \qquad (1.9)$$

CHAPTER 2

UNIVARIATE MODEL BUILDING

We established in Chapter 1 the representation of a stochastic process in autoregressive (AR), moving average (MA) and mixed (ARMA) form. We now study certain characteristics of the different forms for several orders. The necessary restrictions to be placed on each model are indicated. A description of the stages in the modelling procedure is then given.

2.1 AUTOREGRESSIVE PROCESSES

Equation (1.5) represented an autoregressive process with an infinite number of parameters. Assume that the process X_t has zero mean and is represented by

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + a_t \qquad (2.1)$$

which is an autoregressive process of order p [AR(p)]. The series $\{a_t\}$ is a white noise series which is also uncorrelated with all previous X_t's. The set of parameters $\{\phi_1, \phi_2, \ldots, \phi_p\}$ is finite. Equation (2.1) can be written as

$$\phi(B) X_t = a_t$$

where $$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \qquad (2.2)$$

is a polynomial of order p. If we factorize $\phi(B)$ then

$$\phi(B) = \Pi_{j=1}^{p}(1-G_jB) \tag{2.3}$$

where $G_1^{-1},\ldots,G_p^{-1}$ are the different roots of $\phi(B) = 0$.

Expressing X_t in terms of a_t's only we have

$$X_t = \phi^{-1}(B)a_t$$

$$= \Pi_{j=1}^{p}(1-G_jB)^{-1}a_t \tag{2.4}$$

If equation (2.4) is expanded in partial fractions then

$$X_t = \sum_{j=1}^{p} \frac{k_j}{1-G_jB}\, a_t$$

with k_j constants.

Now

$$\frac{k_j}{1-G_jB}\, a_t = k_j \sum_{s=0}^{\infty} G_j^s B^s a_t \quad \text{if} \quad |G_j| < 1 \tag{2.5}$$

Thus

$$X_t = \sum_{j=1}^{p} k_j \sum_{s=0}^{\infty} G_j^s B^s a_t \quad \text{if} \quad |G_j| < 1 \quad \text{for all} \quad j$$

Now

$$E[X_t] = \sum_{j=1}^{p} k_j \sum_{s=0}^{\infty} G_j^s\, E[a_{t-s}]$$

$$= 0 \tag{2.6}$$

and

$$E[X_t^2] = E\left[\sum_{j=1}^{p} k_j \sum_{s=0}^{\infty} G_j^s a_{t-s}\right]\left[\sum_{r=1}^{p} k_r \sum_{\alpha=0}^{\infty} G_j^{\alpha} a_{t-\alpha}\right]$$

$$= \sum_{j=1}^{p}\sum_{r=1}^{p} k_j k_r \left\{\sum_{s=0}^{\infty} G_j^{2s} E[a_{t-s}^2] + \sum_{s\neq\alpha=0}^{\infty} G_j^{s+\alpha} E[a_{t-s}]E[a_{t-\alpha}]\right\}$$

$$= \sum_{j=1}^{p}\sum_{r=1}^{p} k_j k_r \sum_{s=0}^{\infty} G_j^{2s}\, \sigma_a^2 \tag{2.7}$$

The $E[X_t]$ and $\mathrm{var}[X_t]$ are thus independent of time

which are the requirements for a series to be stationary. For the process to be stationary we required that $|G_j| < 1$ for all j which implies that $|G_j^{-1}| > 1$ for all j. Thus the roots of $\phi(B) = 0$ all lie outside the unit circle.

2.1.1 AUTOCORRELATION FUNCTION

If we multiply equation (2.1) by X_{t-k} and take expected values, then

$$E[X_t X_{t-k}] = \phi_1 E[X_{t-1} X_{t-k}] + \ldots + \phi_p E[X_{t-p} X_{t-k}] + E[a_t X_{t-k}].$$

But $E[a_t X_{t-k}] = 0$ if $k > 0$ because a_t is uncorrelated with all previous X_t's. Thus

$$\lambda_k = \phi_1 \lambda_{k-1} + \ldots + \phi_p \lambda_{k-p} \quad \text{if} \quad k > 0 \tag{2.8}$$

and

$$\lambda_0 = \phi_1 \lambda_1 + \ldots + \phi_p \lambda_p + \sigma_a^2$$

Further

$$\lambda_0 [1 - \phi_1 \rho_1 - \ldots - \phi_p \rho_p] = \sigma_a^2$$

so that

$$\lambda_0 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \ldots - \phi_p \rho_p} \tag{2.9}$$

which gives an estimated value of the variance.
Dividing equation (2.8) by λ_0 gives

$$\rho_k = \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p} \quad \text{for} \quad k > 0 \tag{2.10}$$

The general solution for equation (2.10) is

$$\rho_k = A_1 G_1^k + A_2 G_2^k + \ldots + A_p G_p^k \tag{2.11}$$

where G_j^{-1}, $j = 1,\ldots,p$ are the roots of

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p = 0$$

We required that $|G_j| < 1$, $j = 1,\ldots,p$ for the process to be stationary. Two situations can occur in practice:

(i) If G_j is real then $A_j G_j^k$ decays geometrically to zero as k increases which is a damped exponential.

(ii) If the pair G_i, G_j is complex they contribute a term $d^k \sin(2\pi fk+F)$ to ρ_k which follows a damped sine wave as k increases. d is the damping factor, f is the frequency and F is the phase.

The following sets of equations are obtained by substituting $k = 1,2,\ldots,p$ into equation (2.10):

$$\begin{aligned}
\rho_1 &= \phi_1 + \rho_1 \phi_2 + \ldots + \rho_{p-1}\phi_p \\
\rho_2 &= \rho_1 \phi_1 + \phi_2 + \ldots + \rho_{p-2}\phi_p \\
&\vdots \\
\rho_p &= \rho_{p-1}\phi_1 + \rho_{p-2}\phi_2 + \ldots + \phi_p
\end{aligned} \qquad (2.12)$$

These equations are called the Yule-Walker equations

Let

$$\underline{\rho} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \cdot \\ \rho_p \end{bmatrix} \qquad P = \begin{bmatrix} 1 & & \cdots & \rho_{p-1} \\ \rho_1 & 1 & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \rho_{p-1} & & & 1 \end{bmatrix} \qquad \underline{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \cdot \\ \cdot \\ \cdot \\ \theta_p \end{bmatrix}$$

Then equation (2.12) can be written as

$$\underline{\rho} = P\,\underline{\phi}$$

so that

$$\underline{\phi} = P^{-1}\,\underline{\rho} \qquad (2.13)$$

Values for $\underline{\phi}$ is obtained from equation (2.13) where the ρ's in $\underline{\rho}$ are replaced by the estimated autocorrelations

$$r_k = c_k/c_0$$

with

$$c_k = \frac{1}{n}\sum_{t=1}^{n-k}(X_t-\overline{X})(X_{t+k}-\overline{X}) \qquad k = 0,1,\ldots,p$$

and

$$\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$$

The value of ϕ_p in equation (2.13) equals the p^{th} partial autocorrelation Φ_{pp}. This can be proved by taking different values for p and solving the set of equations. This provides an easy way of calculating the partial autocorrelations.

If an AR(p) is the true model then

$$\Phi_{kk} = 0 \quad \text{for} \quad k > p$$

so that the partial autocorrelation function cuts off after

lag p. This is very helpful to find the order of the AR process when unknown as the order is just the lag at which the partial autocorrelation cuts off. When an AR model is fitted to a data series the order will be usually one or two. We now look at an AR(1) and AR(2) process individually using the results established for the general case.

2.1.2 AN AR(1) PROCESS

For the process to be stationary we require that the root of $1-\phi_1 B = 0$ lie outside the unit circle which restricts $|\phi_1| < 1$. The autocorrelation function is

$$\rho_k = \phi_1 \rho_{k-1} \qquad \text{for} \qquad k > 0$$

$$= 1 \qquad \text{for} \qquad k = 0$$

The autocorrelation function decays exponentially to zero. If $\phi_1 < 0$ the decay alternates in sign. An initial estimate for ϕ_1 is r_1, the estimate of ρ_1.

The partial autocorrelation function consists of a non-zero value at lag 1. $\Phi_{kk} = \phi_1 \qquad \text{for} \qquad k = 1$

$$= 0 \qquad \text{for} \qquad k > 1$$

The variance of X_t is

$$\lambda_0 = \frac{\sigma_a^2}{1-\rho_1\phi_1} = \frac{\sigma_a^2}{1-\phi_1^2}$$

2.1.3 AN AR(2) PROCESS

For an AR(2) process to be stationary we require that the roots of $1-\phi_1 B - \phi_2 B^2 = 0$ lie outside the unit circle. Let $\phi(B) = 1-\phi_1 B - \phi_2 B^2 = (1-G_1 B)(1-G_2 B)$. Then the roots of $\phi(B) = 0$ are G_1^{-1} and G_2^{-1}. For stationarity we require that $|G_1| < 1$ and $|G_2| < 1$.

This restricts the parameters ϕ_1 and ϕ_2 to the region

$$\phi_1 + \phi_2 < 1$$

$$-\phi_1 + \phi_2 < 1$$

$$-1 < \phi_2 < 1$$

The autocorrelation function is

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} \qquad \text{for } k > 1$$

with initial values

$$\rho_0 = 1$$

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$

The partial autocorrelation function is

$$\Phi_{11} = \rho_1 = \frac{\phi_1}{1-\phi_2}$$

$$\Phi_{22} = \phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}$$

$$\Phi_{kk} = 0 \qquad \text{for } k > 2$$

The function thus cuts off after lag 2 for an AR(2) process.

The variance is

$$\lambda_0 = \frac{\sigma_a^2}{1-\phi_1\rho_1-\phi_2\rho_2}$$

$$= \frac{\sigma_a^2}{1 - \frac{\phi_1}{1-\phi_2} - \frac{\phi_2(\phi_2-\phi_2^2-\phi_1^2)}{1-\phi_2}}$$

$$= \frac{\sigma_a^2(1-\phi_2)}{(1+\phi_2)[(1-\phi_2)^2-\phi_1^2]}$$

Initial estimates for ϕ_1 and ϕ_2 are

$$\hat{\phi}_2 = \Phi_{22}$$

$$\hat{\phi}_1 = r_1(1-\hat{\phi}_2)$$

2.2 MOVING AVERAGE PROCESSES

Equation (1.4) is the representation of a moving average process with an infinite number of parameters. Assume the process X_t has zero mean. Then

$$X_t = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \qquad (2.14)$$

is a moving average process of order q [MA(q)]. The sequence $\{a_t\}$ is a white noise series. We can write the model as

$$X_t = \theta(B)a_t$$

with

$$\theta(B) = 1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q \qquad (2.15)$$

Now

$$E(X_t) = \theta(B)E(a_t)$$

$$= 0$$

and

$$\lambda_k = E[X_t X_{t-k}]$$

$$= E[(a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \ldots - \theta_q a_{t-k-q})]$$

$$= [-\theta_k + \theta_1\theta_{k+1} + \ldots + \theta_{q-k}\theta_q]\sigma_a^2 \quad \text{if } k = 0,1,\ldots,q$$

$$\text{or } 0 \quad \text{if } k > q \ .$$

The process is stationary as the mean and autocovariance are independent of time and the a_t's are normally distributed. Further, the variance of the process is

$$\lambda_0 = E[X_t^2] = (1+\theta_1^2+\ldots+\theta_q^2)\sigma_a^2$$

2.2.1 AUTOCORRELATION FUNCTION

The autocorrelation function is

$$\rho_k = \lambda_k/\lambda_0$$

$$= \frac{[-\theta_k\theta_1+\theta_1\theta_{k+1}+\ldots+\theta_{q-k}\theta_q]}{1+\theta_1^2+\ldots+\theta_q^2} \quad \text{if } k = 0,1,\ldots,q$$

$$= 0 \quad \text{if } k > q \qquad (2.16)$$

Thus for a MA(q) process the autocorrelation cuts off after lag q, the order of the process. The value of ρ_k is limited for all k. Davies (1974) showed that the maximum value of the k^{th} autocorrelation of a MA(q) process is

$$\rho_k = \cos[\tfrac{\pi}{N+1}] \qquad \text{if } k \text{ divides } q+1$$

$$= \cos[\tfrac{\pi}{N+2}] \qquad \text{if } k \text{ does not divide } q+1$$

2.2.2 INVERTIBILITY CONDITION

We can write $X_t = \theta(B)a_t$ as

$$a_t = \theta^{-1}(B)X_t \tag{2.17}$$

Thus a_t is expressed as an infinite linear function of current and previous X_t's. If we factorize

$$\theta(B) = (1-\theta_1 B - \ldots - \theta_q B^q)$$

then

$$\theta(B) = \Pi_{j=1}^{q}(1-G_j B) \tag{2.18}$$

Substituting equation (2.18) in (2.17) gives

$$a_t = \Pi_{j=1}^{q}(1-G_j B)^{-1}X_t$$

$$= \sum_{j=1}^{q} \frac{k_j}{1-G_j B} X_t \quad \text{where } k_j \text{ are constants}$$

$$= \sum_{j=1}^{q} k_j \sum_{r=0}^{\infty} G_j^r B^r X_t \quad \text{if } |G_j| < 1 \text{ for all } j \tag{2.19}$$

Restricting $|G_j| < 1$ for all j implies that the roots of $\theta(B) = 0$ lie outside the unit circle. The constraints introduced when expressing a MA(q) process as an AR(∞) process is called the invertibility condition. The restrictions are similar to those for stationarity when we express an AR(p) process as a MA(∞) process.

2.2.3 PARTIAL AUTOCORRELATION FUNCTION

The exact expression of the partial autocorrelation function for a MA(q) process is complicated. It is dominated by damped exponentials if the roots are real, damped sine waves if the roots are complex and a mixture of damped exponentials and damped sine waves if the roots are real and complex.

2.2.4 MA(1) PROCESS

A MA(1) process is represented by

$$X_t = a_t - \theta_1 a_{t-1}$$
$$= (1-\theta_1 B)a_t$$

For the process to be invertible we restrict θ_1 to

$$-1 < \theta_1 < 1$$

The autocorrelation function is

$$\rho_k = 1 \qquad \text{if } k = 0$$
$$= -\frac{\theta_1}{1+\theta_1^2} \qquad \text{if } k = 1$$
$$= 0 \qquad \text{if } k > 1$$

The partial autocorrelation function is calculated by writing it in terms of autocorrelations which gives the following general expression

$$\Phi_{kk} = \frac{-\theta_1^k(1-\theta_1^2)}{1-\theta_1^{2(k+1)}} \qquad k = 1,2,\ldots$$

Further $|\Phi_{kk}| < \theta_1^k$ so that it follows a damped exponential decay to zero. If ρ_1 is positive, then θ_1 is negative and the partial autocorrelation alternates in sign. If ρ_1 is negative then θ_1 is positive so that all the partial autocorrelations will be negative. An initial estimate for θ_1 is obtained by solving the quadratic equation

$$\rho_1 = \frac{-\theta_1}{1+\theta_1^2}$$

which is equal to

$$\theta_1^2 - \frac{1}{\rho_1}\theta_1 + 1 = 0$$

The products of the roots are unity. If θ_1 is a root then θ_1^{-1} will be the other root. If θ_1 satisfies the condition for invertibility, then θ_1^{-1} will lie outside the region.

2.2.5 MA(2) PROCESS

The MA(2) process is represented by

$$\begin{aligned} X_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \\ &= (1-\theta_1 B - \theta_2 B^2)a_t \\ &= \theta(B)a_t \end{aligned}$$

The process will be invertible if the roots of the equation

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 = 0 \qquad (2.20)$$

lie outside the unit circle which restricts θ_1 and θ_2 to the region

$$\theta_1 + \theta_2 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$-1 < \theta_2 < 1 \qquad (2.21)$$

The autocorrelation function is

$$\rho_k = \frac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2} \qquad \text{if} \quad k = 1$$

$$= \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} \qquad \text{if} \quad k = 2$$

$$= 0 \qquad \text{if} \quad k > 0 \qquad (2.22)$$

The restrictions on θ_1 and θ_2 imply restrictions on ρ_1 and ρ_2. From equations (2.21) and (2.22) we find that ρ_1 and ρ_2 are limited to the area bounded by

$$\rho_1 + \rho_2 = -\tfrac{1}{2}$$

$$-\rho_1 + \rho_2 = \tfrac{1}{2}$$

$$\rho_1^2 = 4\rho^2(1-\rho^2) \qquad (2.23)$$

The partial autocorrelation function is dominated by damped exponential waves if the roots of equations (2.20) are real, but by damped sine waves if the roots are complex.

2.3 AUTOREGRESSIVE-MOVING AVERAGE PROCESSES

The Autoregressive-Moving Average process [ARMA (p,q)] is represented by

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q} \qquad (2.24)$$

Now

$$(1-\phi_1 B -\ldots-\phi_p B^p)X_t = (1-\theta_1 B-\ldots-\theta_q B^q)a_t$$

thus

$$\phi(B)X_t = \theta(B)a_t$$

We assume that the process has zero mean. If the mean is non-zero then X_t is replaced by $\tilde{X}_t = X_t-\mu$ for all t. The process will be stationary if the roots of $\phi(B) = 0$ lie outside the unit circle. If the process is stationary then we can write it as a $MA(\infty)$ process

$$X_t = c(B)a_t$$

where

$$c(B) = \sum_{j=o}^{\infty} c_j B^j = \theta(B)/\phi(B) \qquad (2.25)$$

The invertibility conditions hold if the roots of $\theta(B) = 0$ all lie outside the unit circle. The equivalent $AR(\infty)$ process is

$$d(B)X_t = a_t$$

where

$$d(B) = \sum_{j=o}^{\infty} d_j B^j = \phi(B)/\theta(B) \qquad (2.26)$$

If we multiply equation (2.24) by X_{t-k} and take expected values, then

$$E[X_t X_{t-k}] = \phi_1 E[X_{t-1} X_{t-k}] + \phi_2 E[X_{t-2} X_{t-k}]+\ldots+\phi_p E[X_{t-p} X_{t-k}]$$

$$+ E[a_t X_{t-k}] - \theta_1 E[a_{t-1} X_{t-k}]-\ldots-\theta_q E[a_{t-q} X_{t-k}] \qquad (2.27)$$

Let

$$\gamma_{\tilde{X}a}(k) = \mathrm{cov}[X_t, a_{t-k}]$$

Then equation (2.27) becomes

$$\lambda_k = \phi_1\lambda_{k-1} + \phi_2\lambda_{k-2} + \ldots + \phi_p\lambda_{k-p} + \gamma_{xa}(k)$$

$$-\theta_1\gamma_{xa}(k-1) - \ldots - \theta_q\gamma_{xa}(k-q) \qquad (2.28)$$

But X_t only depends on the random variables $a_t, a_{t-1},\ldots$ Thus

$$\gamma_{xa}(k) = 0 \quad \text{if} \quad k > 0$$

$$\neq 0 \quad \text{if} \quad k < 0$$

From equation (2.28)

$$\lambda_k = \phi_1\lambda_{k-1} + \phi_2\lambda_{k-2} + \ldots + \phi_p\lambda_{k-p} \qquad (2.29)$$

$$\text{if} \quad k > q$$

Dividing λ_k by λ_0 gives

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \ldots + \phi_p\rho_{k-p} \qquad (2.30)$$

$$\text{if} \quad k > q$$

To find $\phi_1,\ldots,\phi_p$ from equation (2.30) we require the starting values $\rho_q,\rho_{q-1},\ldots,\rho_{q-p+1}$. The initial values $\rho_1,\ldots,\rho_q$ will depend directly on $\phi_1,\ldots,\phi_p$ and $\theta_1,\ldots,\theta_q$ while $\rho_{q+1}, \rho_{q+2},\ldots$ only depends directly on $\phi_1,\ldots,\phi_p$. If $q > p$ then there are $q-p+1$ autocorrelations $\rho_1,\ldots,\rho_{q-p+1}$ which do not follow the pattern while all the other autocorrelations consist of damped exponentials and/or damped sine waves depending on whether the roots are real, complex or both.

The variance using equation (2.28) is

$$\lambda_0 = \phi_1\lambda_1 + \ldots \phi_p\lambda_p + \gamma_{xa}(0) - \theta_1\gamma_{xa}(-1) - \ldots - \theta_q\gamma_{xa}(-q) \qquad (2.31)$$

The exact representation of the partial autocorrelation function is very complicated. It is an infinite decaying series and behaves after p-q lags like a moving average process.

2.3.1 ARMA (1,1) PROCESS

In practice an ARMA model will rarely have more than one autoregressive and one moving average parameter. We consider thus an ARMA (1,1) process

$$X_t = \phi_1 X_{t-1} + a_t - \theta_1 a_{t-1} \qquad (2.32)$$

The process will be stationary if the root of $1 - \phi B = 0$ lies outside the unit circle which implies that

$$|\phi_1| < 1 \qquad (2.33)$$

The invertibility condition will be satisfied if the root of $1 - \theta_1 B = 0$ lies outside the unit circle, implying

$$|\theta_1| < 1 \qquad (2.34)$$

Multiplying equation (2.32) by a_{t-1} and taking expected values gives

$$\begin{aligned}\gamma_{xa}(-1) &= E[X_t a_{t-1}] \\ &= \phi_1 E[X_{t-1} a_{t-1}] + E[a_t a_{t-1}] - E[a_{t-1}^2] \\ &= (\phi_1 - \theta_1)\sigma_a^2\end{aligned}$$

From equation (2.28) we have

$$\lambda_0 = \phi_1\lambda_1 + \sigma_a^2 - \theta_1\gamma_{xa}(-1)$$

$$\lambda_1 = \phi_1\lambda_0 - \theta_1\sigma_a^2$$

so that

$$\lambda_0 = \frac{\sigma_a^2}{1-\phi_1^2}\,[1+\theta_1^2 - 2\theta_1\phi_1]$$

and

$$\lambda_1 = \frac{\sigma_a^2}{1-\phi_1^2}\,(\phi_1-\theta_1)(1-\phi_1\theta_1)$$

The autocorrelation function is

$$\rho_k = \frac{(\phi_1-\theta_1)(1-\phi_1\theta_1)}{1+\theta_1^2-2\phi_1\theta_1} \qquad \text{if} \quad k = 1$$

$$= \phi_1\,\rho_{k-1} \qquad \text{if} \quad k > 1 \qquad (2.35)$$

As the parameter values are unknown we can restrict ρ_1 and ρ_2 to a region that corresponds to $|\phi_1| < 1$ and $|\theta_1| < 1$ so that the process will be stationary and invertible. Using equation (2.35) we find that ρ_1 and ρ_2 are bounded to the area which lies within the region

$$|\rho_2| < |\rho_1|$$

$$\rho_2 > \rho_1(2\rho_1+1) \qquad \text{if} \qquad \rho_1 > 0$$

$$\rho_2 > \rho_1(2\rho_1-1) \qquad \text{if} \qquad \rho_1 < 0 \qquad (2.36)$$

2.4 NONSTATIONARY PROCESSES

In the models previously discussed we restricted them to be stationary. Most economic time series have no fixed mean

and the conditions for weak stationarity will thus not hold. Most series are homogeneous because one part of a series behaves like any other part and only a shift in level has occurred. Another characteristic of economic time series is that the variation is roughly proportional to the level of the series. No exact techniques exist for testing whether a series has a linear trend in mean or in variance. A plot of the series can be very useful. Another form of nonstationary behaviour which is difficult to detect is displayed when parameters change with time. One should therefor only consider processes for which parameters are expected to be constant or change slowly throughout time.

A nonstationary process is known as an Autoregressive Integrated Moving Average process. It is represented by

$$\phi(B)\ (1-B)^d\ X_t = \theta(B)a_t \qquad (2.37)$$

This is equivalent to an ARMA process where d of the parameters are equal to unity. d equals 1 if the process is nonstationary in level and is 2 when it is nonstationary in level and in slope.

The autocorrelation function of a nonstationary series has a unique pattern. Consider the model

$$X_t = X_{t-1} + \theta_0 + a_t \qquad (2.38)$$

If $\theta_0 \neq 0$ then we include a deterministic trend in the model. Assume the series starts at time zero with initial value A. Then

$$X_1 = A + \theta_0 + a_1$$

$$X_2 = A + 2\theta_0 + a_1 + a_2$$

$$\vdots$$

$$X_n = A + n\theta_0 + \sum_1^j a_j$$

The expected value of X_t is

$$E[X_t] = A + t\theta_0 \tag{2.39}$$

which is a linear function of time.

The variance of X_t is

$$\begin{aligned}\lambda_0(t) &= E[X_t - E(X_t)]^2 \\ &= E[\textstyle\sum_1^t a_j]^2 \\ &= \textstyle\sum_1^t E[a_j^2] \\ &= \textstyle\sum_1^t \sigma_a^2 \\ &= t\sigma_a^2\end{aligned}$$

The autocovariance function is

$$\begin{aligned}\lambda_k(t) &= E[X_t - E(X_t)][X_{t-k} - E(X_{t-k})] \\ &= E[\textstyle\sum_{j=1}^t a_j][\sum_{i=1}^{t-k} a_i] \\ &= (t-k)\sigma_a^2\end{aligned}$$

so that the autocorrelation function is

$$\rho_k(t) = \frac{E[X_t - E(X_t)][X_{t-k} - E(X_{t-k})]}{\sqrt{E[X_t - E(X_t)]^2}\ \sqrt{E[X_{t-k} - E(X_{t-k})]^2}}$$

$$= \frac{(t-k)\sigma_a^2}{\sqrt{t\sigma_a^2}\ \sqrt{(t-k)\sigma_a^2}}$$

$$= \sqrt{\frac{t-k}{t}} \qquad (2.40)$$

If t is large in relation to k then

$$\rho_k(t) \simeq 1 \qquad \text{for all } k$$

The series is nonstationary as its mean and variance are increasing with time. In practice we would usually exclude a deterministic trend because it limits the model to follow exactly the same pattern as in the past. A stochastic trend does not have the above limitation.

2.5 THE MODEL BUILDING PROCEDURE

In the previous sections we discussed different kinds of models and the characteristics that they exhibit. In this section we explain how a data series is analysed. We first identify the appropriate model, calculate initial estimates for the parameters included in the model, estimate the model by least squares and then check if the model is adequate.

2.5.1 IDENTIFICATION

The most general model to be fitted is an ARIMA (p,d,q) model. We first have to obtain the level of differencing d and then find the values of p and q. The analyser will first look at a plot of the series X_t against time t. From the plot one will be able to see if the series is stationary

in level and in slope. More accurate conclusions can be made from the sample autocorrelation and the sample partial autocorrelation function. Nonstationarity is indicated if the autocorrelations fail to damp out. The autocorrelations do not have to be extremely large. To remove the nonstationary behaviour of the series we difference the series until it is stationary. We need only to identify the lowest level of differencing for which a stationary series will be apparent as the differences of a stationary series are again stationary and nothing is gained by further differencing.

The necessary degree of differencing d is reached when the sample autocorrelation function of $\nabla^d X_t$ dies out quickly. Overdifferencing will introduce significant partial autocorrelations in the sample partial autocorrelation function. This will lead to the identification of a model of the form

$$(1-B)^d X_t = (1-\theta B)a_t$$

with θ very close to one. This includes a redundant factor which cancels out. Thus the above model is equivalent to

$$(1-B)^{d-1} X_t = a_t$$

The value of d is usually 0,1 or 2. We usually inspect the first 25 estimated autocorrelations and partial autocorrelations of the original and differenced series.

The sample autocorrelation and sample partial autocorrelation function of the differenced series $\nabla^d X_t$ are far more

likely to provide useful information about the process as this information will be damped by the nonstationary behaviour of the series.

We now have to find the values of p and q. As the sample autocorrelations can have large variances and also be highly correlated with each other, it is possible that the sample autocorrelation function will not exhibit the properties of the theoretical autocorrelation function. The sample autocorrelation function can include moderately large values after the theoretical autocorrelation function has damped out. The occurrence of ripples and trends which do not follow a theoretical pattern is also possible.

We have to test the hypothesis that individual autocorrelations and partial autocorrelations are significantly different from zero. Bartlett (1946) showed that the standard error of the estimated autocorrelations is approximately

$$\hat{\sigma}(r_k) \simeq \frac{1}{\sqrt{n}} \left[1 + 2(r_1^2+r_2^2+\ldots+r_q^2)\right]^{\frac{1}{2}} \qquad (2.41)$$

for $k > q$, if we want to test that the autocorrelations are zero after lag q, i.e. the series is described by a MA(q) process. Quenouille (1949) proved that assuming the process is AR(p), the standard error of the sample partial autocorrelation is approximately

$$\hat{\sigma}(\hat{\Phi}_{kk}) \simeq \frac{1}{\sqrt{n}} \qquad \text{for} \quad k > p \qquad (2.42)$$

R.L. Anderson (1942) showed that for moderate n ($n > 50$)

the distribution of the theoretical autocorrelations is approximately Normal. A similar result holds for the theoretical partial autocorrelations.

Combining the above results we can test the hypothesis that ρ_k is zero. The test statistic is

$$\frac{\gamma_k}{\hat{\sigma}(r_k)} \quad \text{which is distributed} \quad N(0,1)$$

The hypothesis will be rejected at 95% confidence level if

$$\left|\frac{r_k}{\hat{\sigma}(r_k)}\right| > 1,96$$

To test the hypothesis $\Phi_{kk} = 0$ we compare

$$\left|\frac{\hat{\Phi}_{kk}}{\hat{\sigma}(\hat{\Phi}_{kk})}\right| \simeq |\sqrt{n}\ \hat{\Phi}_{kk}| \quad \text{against an upper 95\% confidence point}$$

of $N(0,1)$ distribution. If $|\sqrt{n}\ \hat{\Phi}_{kk}| > 1,96$ the hypothesis will be rejected.

We now know which sample autocorrelations and sample partial autocorrelations differ significantly from zero. We can decide whether an autoregressive, moving average or a mixed model should be fitted and what the values of p and q are respectively. We then investigate whether the significant sample autocorrelations do lie within the restricted area for the model identified so that it is stationary and invertible. In Sections 2.1 to 2.3 we mentioned the areas to which the autocorrelations are restricted for an AR model of order 1 and

2, a MA model of order 1 and 2 and an ARMA(1,1) model. If the model is plausible we proceed to obtain initial estimates for the parameters included in the model. The initial estimates are the solutions of the equations where we expressed the sample autocorrelations as a function of the parameters. An easier but less accurate method is to read off values for the parameter estimates from charts produced by Stralkowski (1968). See Box and Jenkins (1970) pages 517-520 for the appropriate charts.

We test whether a deterministic component should be included in the model. Let $W_t = (1-B)^d X_t$ be the stationary series. Then for moderately large samples $\overline{W} = \frac{1}{n} \sum_{t=1}^{n} W_t$ will be normally distributed with mean μ_w and variance σ_w^2. To test the hypothesis that μ_w is zero we compare $|\overline{W}/\sigma_w|$ with an upper 95% confidence point of N(0,1) distribution. We reject the hypothesis if $|\overline{W}/\sigma_w| > 1,96$.

The estimate of σ_w depends on the kind of model fitted to the data. Box and Jenkins (1970) pages 193-195 obtained the following approximate standard errors.

AR(1) $$\sigma_{\overline{w}} = \sqrt{\frac{\hat{\sigma}^2(1+r_1)}{n(1-r_1)}}$$

AR(2) $$\sigma_{\overline{w}} = \sqrt{\frac{\hat{\sigma}^2(1+r_1)(1-2r_1^2+r_2)}{n(1-r_1)(1-r_2)}}$$

MA(1) $$\sigma_{\overline{w}} = \sqrt{\frac{\hat{\sigma}^2(1+2r_1)}{n}}$$

$$\text{MA(2)} \qquad \sigma_{\bar{w}} = \sqrt{\frac{\hat{\sigma}^2(1+2r_1+2r_2)}{n}}$$

$$\text{ARMA(1,1)} \qquad \sigma_{\bar{w}} = \sqrt{\frac{\hat{\sigma}^2}{n}\left(1 + \frac{2r_1^2}{r_1-r_2}\right)} \qquad (2.43)$$

where $\hat{\sigma}^2$ is the estimate of variance of W_t and n is the number of observations in the differenced series.

2.5.2 ESTIMATION

We have to obtain maximum likelihood estimates of the parameters included in the identified model. Suppose we have $N+d$ observations $X_{-d+1},\ldots,X_0, X_1,\ldots,X_n$ and differencing of order d is necessary to induce stationarity. Let $W_t = \nabla^d X_t$ be the stationary series. We fit an ARMA(p,q) model to $W_1, W_2,\ldots,W_n$. The ARMA(p,q) model can be written as

$$a_t = W_t - \phi_1 W_{t-1} - \ldots - \phi_p W_{t-p} + \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q} \qquad (2.44)$$

For simplicity we assume that the mean of W_t is zero.

2.5.2.1 THE CONDITIONAL APPROACH

Suppose we have p values $W_0, W_{-1},\ldots,W_{-p+1}$ and q values $a_0, a_{-1},\ldots,a_{1-q}$ prior to the starting point of the time series. Let W_* and a_* represent these values. We also assume that the a_t's are normally distributed. Given a set of data $W = [W_1,\ldots,W_n]$ then the log likelihood function conditional on the sets W* and a* is

$$L_*(\phi,\theta,\sigma_a) = -n \ln \sigma_a - S_*(\phi,\theta)/2\sigma_a^2 \qquad (2.45)$$

where

$$S_*(\phi,\theta) = \sum_{t=1}^{n}[a_t(\phi,\theta|W_*,a_*,W)]^2 \qquad (2.46)$$

To find the starting values W_* and a_* cause problems. The unconditional expectations of a_t and W_t which are zero can be used. These values are not satisfactory if a root of $\phi(B) = 0$ lies close to the unit circle or when a seasonal model is estimated. In these situations it is better to use the first p observations on a_t as the starting values. The estimation will be done using n-p values, but with long data series the slight loss of information is negligible.

2.5.2.2 THE UNCONDITIONAL APPROACH

The unconditional likelihood function is

$$L(\phi,\theta,\sigma_a) = f(\phi,\theta) - n \ln \sigma_a - \frac{S(\phi,\theta)}{2\sigma_a^2} \qquad (2.47)$$

where $f(\phi,\theta)$ is a function of ϕ and θ.

$$S(\phi,\theta) = \sum_{t=-\infty}^{n}\{E[a_t|\phi,\theta, W]\}^2 \qquad (2.48)$$

If the sample size is moderate or large then the Least Square estimates obtained by minimizing $S(\phi,\theta)$ is very close to the maximum likelihood estimate.

The W_t's are generated by the model

$$\phi(B)W_t = \theta(B)a_t$$

It can also be generated by

$$\phi(F)W_t = \theta(B)e_t$$

where F is the forward shift operator and e_t is a sequence of independent random variables with zero mean and variance $\sigma_e^2 = \sigma_a^2$.

The procedure to calculate the Sum of Squares is

1. Let $E[e_{-j}|\phi,\theta, W] = 0$ for $j = 0,1,2,...$

2. Take $e_{n-j} = 0$ for $j = 0,1,...,p-1$

3. Compute $E[e_j]$, $j = 1,2,...,p$ using

$$E[e_t] = E[W_t]-\phi_1 E[W_{t+1}]-...-\phi_p E[W_{t+p}]+\theta_1 E[e_{t+1}]$$
$$+...+\theta_q E[e_{t+q}]$$

4. Compute $E[W_t]$ backwards until they die out from the equation

$$E[W_t] = E[e_t]-\theta_1 E[e_{t+1}]-...-\theta_q E[e_{t+q}]+ E[W_{t+1}]$$
$$+...+\phi_p E[W_{t+p}]$$

5. Assume $E[W_t]$ dies out at $t = 1-Q$. Take

$$E[a_{-t}] = 0 \text{ for } t \geq Q$$

6. The $E[a_t]$ is then calculated using

$$E[a_t] = E[W_t]-\phi_1 E[W_{t-1}]-...-\phi_p E[W_{t-p}]+\theta_1 E[a_{t-1}]$$
$$+...+\theta_q E[a_{t-q}]$$

7. To cancel the effect caused by taking $e_n = e_{n-1} = e_p = 0$

we do a second iterative cycle. The iteration procedure stops when the values calculated for a_t converge.

8. The unconditional Sum of Squares are given by $\{E[a_t]\}^2$

The maximum likelihood estimates is the parameter estimates which minimizes equation (2.47).

Let $\beta = [\phi_1,\ldots,\phi_p, \theta_1,\ldots,\theta_q]$. Assuming that $S(\beta)$ is quadratic then it can be shown that a 95% confidence region for $S(\beta)$ is given by

$$S(\beta) = S(\hat{\beta})\ [1 + \chi^2_k(0,95)/n] \qquad (2.49)$$

where $\chi^2_k(0,95)$ is the upper 95% confidence point of χ^2_k distribution. k equals p+q the number of parameters estimated.

2.5.2.3 NONLINEAR ESTIMATION OF THE PARAMETERS

The maximum likelihood estimate can be approximated by the least squares estimate minimizing

$$S(\phi,\theta) = \sum_{t=-\infty}^{n}\{E[a_t|\phi,\theta,w]\}^2$$

The lower limit of the summation can be taken as 1-Q for practical purposes. The minimization is much simplified if $E[a_t|\phi,\theta,w]$ is a linear function of the parameters ϕ and θ. This holds for all autoregressive models but not for moving average or ARMA models.

We can approximate $E[a_t|\phi,\theta,W]$ satisfactorily if we expand it as a Taylor series round the initial estimated values $\beta_0 = [\phi_{1,0}; \ldots ; \phi_{p,0}; \theta_{1,0}; \ldots ; \theta_{q,0}]$

Thus

$$E[a_t|\beta,W] = E[a_{t,0}|\beta_0,W] - \sum_{i=1}^{k}(\beta_i-\beta_{i,0})Z_i$$

where

$$Z_i = -\frac{\partial}{\partial\beta_i} E[a_t]\Big|_{\beta=\beta_0} \qquad t = 1-Q,\ldots,0,1,\ldots,n$$

In matrix notation we write the $n+Q$ equations as

$$E[a_0] = Z(\beta-\beta_0) + E[a_t] \qquad (2.50)$$

$\beta-\beta_0$ can now be obtained by Ordinary Least Squares. We use the adjustments as a new guess for the parameter values and repeat the procedure until the estimates converge.

2.5.2.4 PARAMETER REDUNDANCY

The estimation procedure create serious difficulties if we fit a model which includes a redundant factor. That means that we include an AR and a MA factor that exactly cancels out or is very close to cancellation. The situation is avoided when the identified model is parsimonious, that is we use the minimum number of parameters in the model to generate the data series.

2.5.3 DIAGNOSTIC CHECKS

After the model is identified and estimated we apply

diagnostic checks to investigate if the fitted model is adequate, and if not, the way in which it can be modified.

2.5.3.1 OVERFITTING

We fit a more elaborate model than the fitted model we believe to be correct by adding an AR or a MA parameter to the model. We do not add both an AR and a MA parameter as this will lead to parameter redundancy which will cause difficulties in the estimation. By overfitting we test the hypothesis that the added parameter is zero. If the 95% confidence interval for the parameter estimate does not include zero, then the added parameter is significant. The autocorrelation function of the stationary series will give an indication whether to add an AR or a MA parameter.

2.5.3.2 AUTOCORRELATIONS IN THE RESIDUALS

We have to test whether autocorrelation is present in the residuals obtained after fitting an ARMA model to the stationary series. R.L. Anderson (1942) has shown that on the assumption that the form of the fitted model is correct and that the true parameter values known, the autocorrelations $r_k(a)$ will be uncorrelated and approximately Normal with zero mean and variance n^{-1} distributed. As the true parameter values are usually unknown, the standard error $(n^{-\frac{1}{2}})$ will underestimate significant departures from zero at low lags but can be employed at moderate and high lags.

Box and Pierce (1970) showed that for any ARIMA model which is appropriate the statistic $Q = n\sum_{k=1}^{h} \hat{r}_k^2(a)$ is distributed as χ^2_{h-p-q} when h is sufficiently large. An inappropriate model is indicated by a high Q-value. The Q-statistic only indicates whether the autocorrelations are generally too high, but it is not a very powerful tool for detecting specific departures from white noise.

2.5.3.3 STABLE PARAMETER ESTIMATES

The analyser can also check if the parameter estimates are time invariant. The data series is divided into two parts and a model is fitted to each part separately. We then test if the difference between the parameter estimates is significant. Let $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ be the estimated values of a parameter. The difference is significant at the 95% confidence level if

$$\left| \frac{\hat{\theta}^{(1)} - \hat{\theta}^{(2)}}{\sqrt{\frac{\hat{\theta}^{(1)}(1-\hat{\theta}^{(1)})}{n_1} + \frac{\hat{\theta}^{(2)}(1-\hat{\theta}^{(2)})}{n_2}}} \right| > 1{,}96$$

where 1,96 is the 95% upper confidence point of a N(0,1) distribution. n_1 and n_2 are the number of observations in each part respectively. If the difference is significant the series should be analysed in two separate parts.

2.6 FORECASTING

Suppose that the ARIMA(p,d,q) model explains the generation of the time series. Then

$$\phi(B)\nabla^d X_t = \theta(B)a_t \qquad (2.51)$$

The ARIMA model can be written as a MA(∞) process

$$X_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

$$= \psi(B)a_t \qquad (2.52)$$

where

$$\psi(B)\phi(B)\nabla^d = \theta(B) \qquad (2.53)$$

Let $\hat{X}_t(\ell)$ be a ℓ-step ahead forecast from origin t of $X_{t+\ell}$. $\hat{X}_t(\ell)$ is a linear function of current and previous observations $X_t, X_{t-1}, \ldots$ and current and previous shocks $a_t, a_{t-1}, \ldots$.

Suppose $\hat{X}_t(\ell)$ is the best forecast of $X_{t+\ell}$ where

$$\hat{X}_t(\ell) = \psi^*_\ell a_t + \psi^*_{\ell+1} a_{t-1} + \psi^*_{\ell+2} a_{t-2} + \ldots \qquad (2.54)$$

The mean square forecast error is

$$E[X_{t+\ell} - \hat{X}_t(\ell)]^2 = E\left[\sum_{j=0}^{\infty}\psi_j a_{t+\ell-j} - \sum_{j=0}^{\infty}\psi^*_{\ell+j} a_{t-j}\right]^2$$

$$= E\left[\sum_{j=0}^{\ell-1}\psi_j a_{t+\ell-j}\right]^2 + E\left[\sum_{j=0}^{\infty}(\psi_{\ell+j} - \psi^*_{\ell+j}) a_{t-j}\right]^2$$

$$= \sum_{j=0}^{\ell-1}\psi_j^2\sigma_a^2 + \sum_{j=0}^{\infty}(\psi_{\ell+j} - \psi^*_{\ell+j})\sigma_a^2 \qquad (2.55)$$

The mean squared forecast error is minimized by setting $\psi^*_{\ell+j} = \psi_{\ell+j}$.

Then

$$X_{t+\ell} = (a_{t+\ell} + \psi_1 a_{t+\ell-1} + \ldots + \psi_\ell a_{t+\ell-1}) + (\psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \ldots)$$

$$= e_t(\ell) + \hat{X}_t(\ell) \qquad (2.56)$$

$e_t(\ell)$ is the error of the ℓ-step ahead forecast $\hat{X}_t(\ell)$.

Let

$$E_t[X_{t+\ell}] = E[X_{t+\ell}|X_t, X_{t-1}, \ldots]$$

Then

$$\hat{X}_t(\ell) = \psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \ldots$$

$$= E_t[X_{t+\ell}] \qquad (2.57)$$

The minimum mean square forecast error for lead time ℓ from origin t is the conditional expectation of $X_{t+\ell}$ at time t. Further

$$e_t(\ell) = a_{t+\ell} + \psi_1 a_{t+\ell-1} + \ldots + \psi_{\ell-1} a_{t+1}$$

Thus

$$E[e_t(\ell)] = 0 \qquad (2.58)$$

The forecast is unbiased. The variance of the forecast error is

$$\text{var}[e_t(\ell)] = E\left[\sum_{j=0}^{\ell-1} \psi_j a_{t+\ell-j}\right]^2$$

$$= \sum_{j=0}^{\ell-1} \psi_j^2 \, \sigma_a^2 \qquad (2.59)$$

The one-step ahead forecast error is

$$e_t(1) = X_{t+1} - \hat{X}_t(1)$$

$$= a_{t+1} \qquad (2.60)$$

which is the residual series. Thus the one-step ahead forecasts must be uncorrelated. The forecasts for ℓ periods ahead ($\ell > 1$) will be correlated as it follows a MA(ℓ-1) process.

Let $\Psi(B) = \phi(B)(1-B)^d$ in equation (2.51). Then we can write the model as

$$X_{t+\ell} = \Psi_1 X_{t+\ell-1} + \ldots + \Psi_{p+d} X_{t+\ell-p-d} + a_{t+\ell}$$
$$- \theta_1 a_{t+\ell-1} - \ldots - \theta_q a_{t+\ell-q}$$

Further

$$\hat{X}_t(\ell) = E_t[X_{t+\ell}]$$
$$= \Psi_1 E_t[X_{t+\ell-1}] + \ldots + \Psi_{p+d} E_t[X_{t+\ell-p-d}] + E_t[a_{t+\ell}]$$
$$- \theta_1 E_t[a_{t+\ell-1}] - \ldots - \theta_q E_t[a_{t+\ell-q}] \qquad (2.61)$$

where

$$E_t[X_{t-j}] = X_{t-j} \qquad j = 0,1,2,\ldots$$
$$E_t[X_{t+j}] = \hat{X}_t(j) \qquad j = 1,2,\ldots$$
$$E_t[a_{t-j}] = a_{t-j} = X_{t-j} - \hat{X}_{t-j-1}(1) \qquad j = 0,1,2,\ldots$$
$$E_t[a_{t+j}] = 0 \qquad j = 1,2,\ldots$$

The forecasts can be generated from equation (2.61)

2.6.1 CONFIDENCE INTERVAL FOR FORECASTS

The ψ-weights can be obtained by equating coefficients in equation (2.53), it is

$$\psi(B)\phi(B)(1-B)^d = \theta(B)$$

They are required in the calculation of $\text{var}[e_t(\ell)]$ in equation (2.59). If we assume that the a_t's are normally distributed, then $X_{t+\ell} | X_t, X_{t-1}, \ldots$ is distributed Normal with mean $\hat{X}_t(\ell)$ and variance $\sum_{j=0}^{\ell-1} \psi_j^2 \sigma_a^2$. A 95% confidence

interval for $X_{t+\ell}$ is

$$\hat{X}_t(\ell) - 1{,}96\sqrt{\sum_{j=0}^{\ell-1}\psi_j^2}\ S_a \leq X_{t+\ell} \leq \hat{X}_t(\ell) + 1{,}96\sqrt{\sum_{j=0}^{\ell-1}\psi_j^2}\ S_a \qquad (2.62)$$

where S_a is the estimate of σ_a.

2.7 SEASONAL MODELS

We excluded seasonal behavioural patterns from the models discussed in the previous sections. It is necessary to distinguish between an additive and a multiplicative seasonal pattern. The pattern is additive if the seasonal effect remains constant throughout time but it is multiplicative if it increases with time.

The additive seasonal effect can be removed by differencing the data series using the transformation $\nabla_s = 1-B^s$ where s is the period of the seasonal effect. Thus if the data displays a linear trend and a seasonal effect then $\nabla\nabla_s$ will reduce it to stationarity.

The removal of a multiplicative seasonal effect can be done in two ways. If the series has a linear trend and a multiplicative seasonal pattern, then differencing the series by $\nabla_s^2 = (1-B^s)^2$ will reduce it to stationarity.

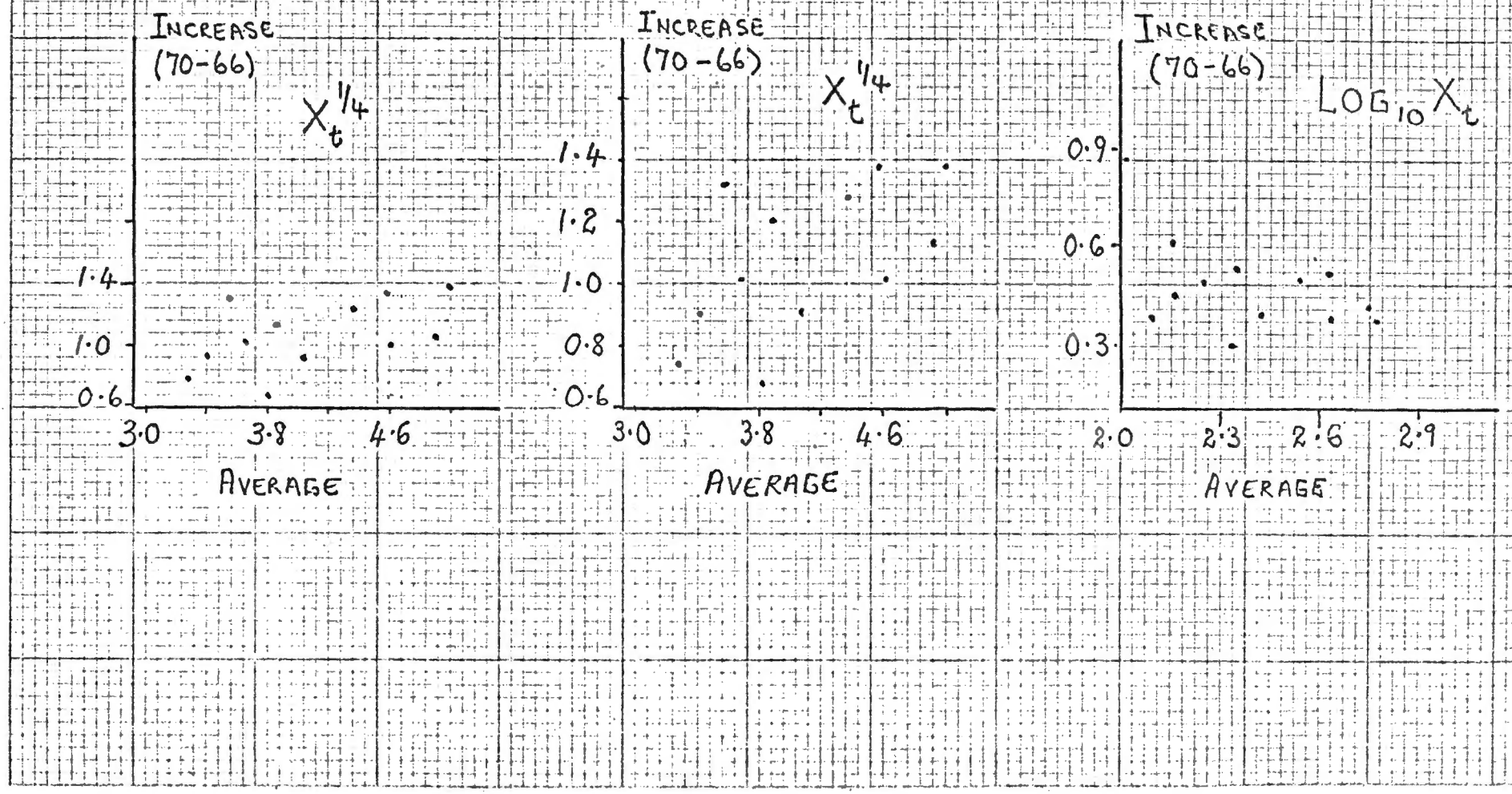
Another way of removing a multiplicative seasonal pattern is by transforming the data series so that the effect becomes additive. The exact transformation can be obtained by

applying the likelihood technique of Box and Cox (1964). An approximate method suggested by Chatfield and Prothero (1973) is to plot the transformed data values at time k, s+k, 2s + k,... where k = 1,2,...,s. We obtain s different graphs. We accept the seasonal pattern as reasonably stable when the different trend lines are roughly linear and parallel. Another approximate method used by Box and Jenkins (1973) suggested the plot of the difference between the value of k in the last and in the first year against the average value of k over the whole period. [k = 1,...,12 assuming monthly data and a yearly seasonal effect.] The transformation is satisfactory if the differences are independent of the size of the data values.

Chatfield and Prothero (1973) felt that the method of Box and Jenkins did not estimate the trend efficiently. In their analysis they found that the $\log_{10}X$ transformation was applicable and not $X^{\frac{1}{4}}$ using the last year and the second years data values while Box and Jenkins showed that $X^{\frac{1}{4}}$ was the best transformation in comparison with X, $X^{\frac{1}{2}}$ and $\log_{10}X$. X is the original data series.

One should use the same scale sizes on both axes when the above mentioned points are plotted. However, Chatfield and Prothero used different scales for $X^{\frac{1}{4}}$ but the same scale for $\log_{10}X$ for their plots. Thus they chose the $\log_{10}X$ transformation which was an incorrect conclusion. The different plots are given in Figure 2.1.

FIGURE 2.1


INCREASE (70-66)
$X_t^{1/4}$
1·4
1·0
0·6
3·0
3·8
4·6
AVERAGE
INCREASE (70-66)
$X_t^{1/4}$
1·4
1·2
1·0
0·8
0·6
3·0
3·8
4·6
AVERAGE
INCREASE (70-66)
$LOG_{10} X_t$
0·9
0·6
0·3
2·0
2·3
2·6
2·9
AVERAGE

The use of transformations creates a large amount of additional analysis. The author therefore preferred to use the difference factor ∇_S^2 instead. The stationary series obtained after differencing can be represented by an ARMA model which may include AR or MA parameters of order S. The approach is very similar to model identification and estimation excluding seasonality.

C H A P T E R 3

TRANSFER FUNCTION MODEL BUILDING

3.1 INTRODUCTION

In Chapter 2 the theory of building a model on present and past data values was discussed. No other information was taken into account in the construction of the model. In economic theory the variation of a variable is usually explained by explanatory variables. In this chapter the response of a variable Y_t to changes in an explanatory variable X_t will be investigated. The response can be delayed and Y_t may return to an equilibrium situation and is then known to be dynamic. The model describing the relationship between Y_t and X_t is called a transfer function model.

A high correlation between variables does not necessarily infer that the variables are causaly related as both may be linked to a common third factor. Correlation is only a measure of a linear relationship so that variables may be functionally related yet uncorrelated.

3.2 CAUSALITY

Granger (1969) gave the following definition of causality.

"A variable X_t causes another variable Y_t with respect to a given information set including both X_t and Y_t if present Y_t can be better predicted by using past values of X_t than by not using them and all other available information is used in either case." The system is not required to be linear.

Let A_t; $t = 0, \pm 1, \pm 2, \ldots$ be an information set that includes at least $\{(X_t, Y_t)\}$. Let $\bar{A}_t = \{A_s : s < t\}$ and $\tilde{A}_t = \{A_s : s \leq t\}$. $\bar{X}_t$, $\tilde{X}_t$, $\bar{Y}_t$ and $\tilde{Y}_t$ are the corresponding definitions for X_t and Y_t respectively. Let $P_t(Y|B)$ denote the mean square error single step predictor of Y_t given an information set B and $\sigma^2(Y|B)$ the mean square error. Variable X_t causes variables Y_t if

$$\sigma^2(Y_t|\bar{A}_t) < \sigma^2(Y_t|\bar{A}_t - \bar{X}_t) \tag{3.1}$$

Variable X_t causes variable Y_t instantaneously if

$$\sigma^2(Y_t|A_t, \tilde{X}_t) < \sigma^2(Y_t|A_t) \tag{3.2}$$

Feedback between X_t and Y_t occurs if X_t causes Y_t and Y_t causes X_t.

We do not require that X_t and Y_t are covariance stationary. Assume there exists transformations so that the transformed series $x_t = T_x X_t$ and $y_t = T_y Y_t$ are covariance stationary. T_x and T_y are usually of the form $\nabla^d \nabla_s^D$ where d and D equal 0, 1 or 2. A causality event will be true for (X_t, Y_t) if and only if it is true for (x_t, y_t).

Consider the general case where feedback exists between x_t and y_t. The model representing (x_t, y_t) is

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \psi_{11}(B) & \psi_{12}(B) \\ \psi_{21}(B) & \psi_{22}(B) \end{bmatrix} \begin{bmatrix} a_t \\ b_t \end{bmatrix} \tag{3.3}$$

a_t and b_t are white noise series. We can write equation (3.3) as

$$\begin{bmatrix} \Pi_{11}(B) & \Pi_{12}(B) \\ \Pi_{21}(B) & \Pi_{22}(B) \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_t \\ b_t \end{bmatrix} \tag{3.4}$$

X_t causes Y_t only if $\psi_{12}(B) = 0$ so that

$$x_t = \psi_{11}(B)\, a_t \tag{3.5}$$

$$\begin{aligned} y_t &= \psi_{21}(B)\, a_t + \psi_{22}(B)\, b_t \\ &= V(B)\, x_t + \psi_{22}(B)\, b_t \end{aligned} \tag{3.6}$$

with $V(B) = \psi_{11}^{-1}(B)\psi_{21}(B)$

Equation (3.5) is the univariate model of x_t in MA(∞) form while equation (3.6) is called the dynamic regression model.

3.3 CROSSCOVARIANCE AND CROSSCORRELATION FUNCTIONS

The crosscovariance between x_t and y_{t+k} is

$$\gamma_{xy}(k) = E[(x_t - \mu_x)(y_{t+k} - \mu_y)] \qquad k = 0, 1, \ldots$$

Further

8. Unidirectional causality from X to Y	$\rho_{xy}(k) \neq 0$ for some $k > 0$ and $\rho_{xy}(k) = 0$ for all $k < 0$ or all $k \leq 0$
9. X and Y are only related instantaneously (if at all)	$\rho_{xy}(k) = 0$ for all $k \neq 0$
10. X and Y are related instantaneously and in no other way	$\rho_{xy}(k) = 0$ for all $k \neq 0$ $\rho_{xy}(0) \neq 0$
11. X and Y are independent	$\rho_{xy}(k) = 0$ for all k

In practice the crosscorrelation coefficients are unknown and each coefficient is estimated by

$$r_{xy}(k) = \frac{C_{xy}(k)}{S_x S_y} \qquad k = 0, \pm 1, \pm 2, \ldots \qquad (3.8)$$

where $C_{xy}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y})$ if $k = 0,1,2,\ldots$

$$= \frac{1}{n} \sum_{t=1}^{n+k} (x_t - \bar{x})(y_{t-k} - \bar{y}) \quad \text{if} \quad k = 0, -1, -2, \ldots$$

From the estimates of the crosscorrelation function we can identify the direction of causality when it exists or when a feedback situation occurs.

3.4 IMPULSE RESPONSE AND STEP RESPONSE FUNCTIONS

Consider equation (3.6)

$$y_t = V(B)x_t + \psi_{22}(B)a_t$$

where

$$V(B) = V_0 + V_1B + V_2B^2 + \ldots \qquad (3.9)$$
$$= \psi_{21}(B)\psi_{11}^{-1}(B)$$

$V(B)$ is called the impulse response function and $V_0, V_1, \ldots$ the impulse response weights. The impulse response function consists of an infinite number of impulse response weights and is not a very practical form to use.

Let

$$SR_j = \sum_{k=0}^{j} V_k$$

and

$$SR(B) = \sum_{j=0}^{\infty} SR_jB^j \qquad (3.10)$$

The values SR_j are the step response weights and $SR(B)$ is the step response function.

Expressing equation (3.6) in a finite form gives

$$\delta(B)y_t = \omega(B)B^bx_t + N_t \qquad (3.11)$$

where

N_t is the generated noise series.

and $\delta(B) = 1 - \delta_1B - \ldots - \delta_rB^r$

$\omega(B) = \omega_0 - \omega_1B - \ldots - \omega_sB^s$

Further

$$\delta(B)V(B) = \omega(B)B^b \qquad (3.12)$$

Equation (3.11) is a transfer function model of order (r,s,b). The generated noise series can be represented by an ARMA(p,q) model

$$\phi(B)N_t = \theta(B)a_t$$

so that equation (3.11) becomes

$$\delta(B)y_t = \omega(B)B^b x_t + \frac{\theta(B)a_t}{\phi(B)} \qquad (3.13)$$

Equation (3.12) enables us to find parameter estimates for $\delta_1,\ldots,\delta_r$; $\omega_1,\ldots,\omega_s$ and b. It equals

$$(1-\delta_1 B - \ldots - \delta_r B^r)(V_0+V_1 B + V_2 B^s + \ldots)$$
$$= (\omega_0-\omega_1 B - \ldots - \omega_s B^s)B^b$$

Equating coefficients for the different powers of B gives

$$\begin{aligned}
V_j &= 0 && \text{for } j < b \\
V_j &= \delta_1 V_{j-1}+\ldots+\delta_r V_{j-r} + \omega_0 && \text{for } j = b \\
V_j &= \delta_1 V_{j-1}+\ldots+\delta_r V_{j-r} - \omega_r && \text{for } j = b+1,\ldots,b+s \\
V_j &= \delta_1 V_{j-1}+\ldots+\delta_r V_{j-r} && \text{for } j > b+s
\end{aligned} \qquad (3.14)$$

The impulse response function consists of

1. b zero values $V_0, V_1,\ldots,V_{b-1}$.
2. s-r+1 values $V_b, V_{b+1},\ldots,V_{b+s-r}$ which do not follow any fixed pattern.
3. V_j's for $j > b+s-r+1$ which follows the pattern dictated by the r^{th} order difference equation. The starting values for the difference equation are $V_{b+s}, V_{b+s-1},\ldots, V_{b+s-r+1}$

In practice the values of r and s are usually 0, 1 or 2 so that we have to consider 9 different functions. The shape of the impulse response and step response functions are

discussed in detail in Box and Jenkins (1970) pages 348-353.

3.5 IDENTIFICATION

The identification stage of the model building procedure can be done in two ways. The approach of Box and Jenkins (1970) consists of applying the same prewhitening transformations to both data series. Prewhitening of the input series is necessary to reduce spurious crosscorrelations which is caused by the presence of autocorrelation in the leading series. The advantage of the above approach is that is is very easy to rewrite the transfer function model in terms of the original data series X_t and Y_t which is useful for forecasting purposes. A disadvantage is that the univariate model for Y_t can be totally different from the model for X_t and thus prewhitening Y_t by the same transformation as X_t will leave significant autocorrelation in the transformed series. This will cause crosscorrelations to be significant which actually are insignificant.

Let x_t be a stationary series which was obtained by differencing X_t. x_t can be represented by the model

$$\phi_x(B)x_t = \theta_x(B)\alpha_t$$

where α_t is a white noise series. We can write α_t in terms of the x_t's. Thus

$$\alpha_t = \phi_x(B)\theta_x^{-1}(B)x_t$$

We now apply $\phi_x(B)\theta_x^{-1}(B)$ to the dependent series y_t which gives us

$$\beta_t = \phi_x(B)\theta_x^{-1}(B)y_t$$

Substituting β_t and α_t for y_t and x_t in equation (3.6) respectively, we obtain

$$\beta_t = V(B)\alpha_t + \varepsilon_t \qquad (3.15)$$

with

$$\varepsilon_t = \phi_x(B)\theta_x^{-1}(B)\psi_{22}(B)a_t$$

If we multiply equation (3.15) by α_{t-k} and take expected values, then

$$E[\beta_t\alpha_{t-k}] = V(B)\ E[\alpha_t\alpha_{t-k}] + E[\varepsilon_t\alpha_{t-k}]$$

so that

$$\gamma_{\alpha\beta}(k) = V_k\ \sigma_\alpha^2$$

Thus

$$V_k = \frac{\gamma_{\alpha\beta}(k)}{\sigma_\alpha^2}$$

But

$$\rho_{\alpha\beta}(k) = \frac{\gamma_{\alpha\beta}(k)}{\sigma_\alpha\sigma_\beta}$$

so that

$$V_k = \rho_{\alpha\beta}(k)\ \frac{\sigma_\beta}{\sigma_\alpha} \qquad \text{for } k = 0,1,2,... \qquad (3.16)$$

The crosscorrelation coefficients are thus proportional to the impulse response weights. Replacing $\rho_{\alpha\beta}(k)$ by its estimate $r_{\alpha\beta}(k)$ and σ_β and σ_α by S_β and S_α respectively we can calculate the impulse response weights from equation (3.16).

Haugh (1972) followed a slightly different approach to Box and Jenkins. To minimize the occurrence of spurious crosscorrelations each series is prewhitened individually. Spurious crosscorrelations can be visually traced if the presence of autocorrelation increases the crosscorrelation coefficient but are impossible to find when it reduces the crosscorrelation coefficient. A disadvantage of Haugh's method is that writing the dynamic model in terms of the original data series will usually lead to an overparametrized model unless the univariate models only have a few coefficients.

Let the univariate models built for the stationary series be

$$\phi_x(B)x_t = \theta_x(B)\alpha_t$$

and $$\phi_y(B)y_t = \theta_y(B)\beta_t$$

where α_t and β_t are white noise series.

If we crosscorrelate α_t and β_t and α_t causes β_t we obtain the dynamic model

$$\beta_t = V(B)\alpha_t + \varepsilon_t \qquad (3.18)$$

where

ε_t is the generated noise series.

The further analysis of the model is exactly the same as for the method by Box and Jenkins.

To test whether a crosscorrelation is significantly different from zero we compare each individual value against its

standard error. As we used a white noise series as a leading series the

$$\text{var}[r_{\alpha\beta}(k)] \simeq (n-k)^{-1} \qquad (3.19)$$

and $\rho[r_{\alpha\beta}(k), r_{\alpha\beta}(k+\ell)] \simeq \rho_{\alpha\alpha}(\ell)$

We accept a crosscorrelation as zero if

$$|\sqrt{n-k}\ r_{\alpha\beta}(k)| < 1{,}96$$

where 1,96 is the upper 95% confidence point of the N(0,1) distribution. An impulse response weight will be significant if the corresponding crosscorrelation coefficient is significant as it is proportional to it.

The value b is the number of zero impulse response weights $V_0, V_1, \ldots, V_{b-1}$. r and s are obtained by finding the theoretical model that agrees the closest with the impulse and step response functions of the data series. Having obtained b, r and s we solve the set of equations (3.14).

We can test if there is an overall significant effect of causality. Under the null hypothesis of independence the statistic

$$Q^* = n^2 \sum_{i=-m}^{m} (N-|i|)^{-1} \hat{r}^2_{\alpha\beta}(k) \qquad (3.20)$$

is asymptotically distributed as χ^2_{2m+1} distribution. We reject the null hypothesis in favour of the alternative that X causes Y if

$$Q^* > \chi^2_{2m+1}(0{,}95)$$

where $\chi^2_{2m+1}(0,95)$ is an upper 95% confidence point of χ^2_{2m+1}.

To identify the generated noise model, we require the autocorrelation function and the partial autocorrelation function. Consider equation (3.15)

$$\beta_t = V(B)\alpha_t + \varepsilon_t$$

$$= u_t + \varepsilon_t$$

On the assumption that u_t and ξ_t is independent we have

$$\gamma_{\beta\beta}(k) = \gamma_{uu}(k) + \gamma_{\varepsilon\varepsilon}(k)$$

and $$\gamma_{uu}(k) = \sigma_\alpha^2 \sum_{j=0}^{\infty} V_j V_{j+k}$$

$$= \frac{1}{\sigma_\alpha^2} \sum_{j=0}^{\infty} \gamma_{\alpha\beta}(j)\gamma_{\alpha\beta}(j+k)$$

Thus $$\gamma_{\varepsilon\varepsilon}(k) = \gamma_{\beta\beta}(k) - \gamma_{uu}(k)$$

$$= \gamma_{\beta\beta}(k) - 1/\sigma_\alpha^2 \sum_{j=0}^{\infty} \gamma_{\alpha\beta}(j)\gamma_{\alpha\beta}(j+k)$$

and $$\gamma_{\varepsilon\varepsilon}(0) = \gamma_{\beta\beta}(0) - 1/\sigma_\alpha^2 \sum_{j=0}^{\infty} \gamma_{\alpha\beta}^2(j)$$

Then $$\rho_{\varepsilon\varepsilon}(k) = \frac{\rho_{\beta\beta}(k) - \sum_{j=0}^{\infty} \rho_{\alpha\beta}(j)\rho_{\alpha\beta}(j+k)}{1 - \sum_{j=0}^{\infty} \rho_{\alpha\beta}^2(j)} \qquad (3.21)$$

The values for $\rho_{\varepsilon\varepsilon}(k)$ are obtained by substituting $r_{\beta\beta}(j)$ and $r_{\alpha\beta}(j)$ for the crosscorrelations in equation (3.21)

The partial autocorrelations are the solutions of the Yule-Walker equations using the estimated autocorrelations.

To limit the possibility of identifying an unnecessarily

complicated model, a simple model should be used first and one should try to simplify it if possible. A more elaborate model will be used if required. Factors in the coefficients of α_t, β_t and ξ_t that are nearly equal should be excluded as it will cause unstable least squares estimates of the parameters.

3.6 ESTIMATION

Assume that we have n observations over which the transfer function relationship

$$\delta(B)\beta_t = \omega(B)B^b\alpha_t + \frac{\theta(B)a_t}{\phi(B)}$$

is estimated. If the starting values β_0, α_0 and a_0 were available prior to the starting point of the data series, then given the data series we could calculate

$$a_t = a_t(b,\delta,\omega,\phi,\theta|\beta_0, \alpha_0, a_0)$$

Assuming that the a_t's are Normally distributed the maximum likelihood estimate can be approximated by minimizing

$$S_0(b,\delta,\omega,\phi,\theta) = \sum_{t=1}^{n} a_t^2(b,\delta,\omega,\phi,\theta|\beta_0,\alpha_0, a_0)$$

We first calculate the β_t's then we find ξ_t's using equation (3.15) and then the a_t's are calculated from the ξ_t's. We require $u = \max[r,\delta+b]$ starting values for the calculation of β_t ; $t = u+1,\ldots,n$. ξ_t will be calculated from $u+1$ onwards. If the unknown a_t's are set equal to their conditional expectations which are zero then the a_t's are calculated from a_{u+p+1} onwards. Thus

$$S_o(b,\delta,\omega,\phi,\theta) = \sum_{t=u+p+1}^{n} a_t^2(b,\delta,\omega,\phi,\theta|\alpha_0,\beta_0,a_0) \qquad (3.22)$$

We can also use a nonlinear least squares algorithm for the estimation. The results will be satisfactory if the sum of squares is roughly quadratic but difficulties occur when parameters are highly correlated or if estimates are close to a boundary limit.

The value b is taken fixed. Let $B_o = (\delta_{1,o},\ldots,\delta_{r,o}; \omega_{0,o},\omega_{1,o},\ldots,\omega_{s,o}; \phi_{1,o},\ldots,\phi_{p,o}; \theta_{1,o},\ldots,\theta_{q,o})$ be the estimates of the parameters and $a_{t,o}$ the corresponding residual value

Let
$$d_{i,t}^{(\delta)} = -\left.\frac{\partial a_i}{\partial \delta_i}\right|_{B_o}$$

$$d_{j,t}^{(\omega)} = -\left.\frac{\partial a_t}{\partial \omega_j}\right|_{B_o}$$

$$d_{g,t}^{(\phi)} = -\left.\frac{\partial a_t}{\partial \phi_g}\right|_{B_o}$$

$$d_{h,t}^{(\theta)} = -\left.\frac{\partial a_t}{\partial \theta_h}\right|_{B_o}$$

We then expand a_t about B_o in a Taylor series

$$a_{t,o} \simeq \sum_{i=1}^{r}(\delta_i-\delta_{i,o})d_{i,t}^{(\delta)} + \sum_{j=o}^{s}(\omega_j-\omega_{j,o})d_{j,t}^{(\omega)}$$

$$+ \sum_{g=1}^{p}(\phi_g-\phi_{g,o})d_{g,t}^{(\phi)} + \sum_{h=o}^{q}(\theta_j-\theta_{j,o})d_{h,t}^{(\theta)} + a_t \qquad (3.23)$$

If t takes the values $1,\ldots,n$ then we can write it in

matrix notation

$$\underline{a}_o = X(B-B_o) + \underline{a}_t \qquad (3.24)$$

where $\underline{a}_o$ and $\underline{a}_t$ are $n\times 1$ vectors

B and B_o are $(r+s+p+q)\times 1$ vectors

X is $n\times(r+s+p+q)$ matrix.

The equation can then be solved by Ordinary Least Squares. We add the solutions to B_o to give a second set of guesses and the estimation procedure is repeated until the parameter estimates converge.

We assumed that b is fixed. If b has to be estimated then the estimation is done for various values of b. We choose the value of b that minimizes the Sum of squares.

3.7 DIAGNOSTIC CHECKS

We employ diagnostic checks to test the adequacy of the model. Extra coefficients can be added to the model and we test whether they are statistically significantly different from zero.

A second check involves the examining of the autocorrelations of the residuals obtained from the fitted model. Let $\beta_t = V(B)\alpha_t + \psi(B)a_t$ be the model that was fitted to the data. Assume that the transfer function specification is correct but that the noise model is incorrect. Let $\psi_0(B)$ be the correct form of the noise model. Then

$$\beta_t = V(B)\alpha_t + \psi_0(B)a_{ot}$$

so that the residuals are

$$a_{ot} = \psi_0^{-1}(B)\psi(B)a_t$$

The residuals are uncorrelated with α_t's but are autocorrelated with itself. Under the null hypothesis that a_t is a white noise series the autocorrelations will be distributed independently with zero mean and standard deviation $n^{-\frac{1}{2}}$. When we employ estimates of the parameters the variance of an autocorrelation at a low lag can be less than $n^{-\frac{1}{2}}$ and the autocorrelations can be highly correlated. We can thus underestimate the significance of an autocorrelation. Box and Jenkins (1970) suggested the following overall check which is not affected by the above distributional effects. If the functional form of a model is adequate then the statistic

$$Q = n \sum_{k=1}^{M} r^2_{\hat{a}\hat{a}}(k) \qquad (3.25)$$

is approximately distributed as χ^2_{k-p-q} distribution where $p+q$ is the number of parameters in the noise model. M must be sufficiently large so that the weights ψ_j, $j > M$ are negligible.

The third check is that the prewhitened α_t and the residuals a_t are mutually stochastically uncorrelated. Inadequacy in the transfer function specification is indicated by a pattern of markedly nonzero cross correlations. If $V_0(B)$ is the correct specification of the transfer function model then

$$\beta_t = V_0(B)\alpha_t + \psi(B)a_{ot}$$

so that

$$a_{ot} = \psi^{-1}(B)[V(B)-V_0(B)]\alpha_t$$

The crosscorrelation analysis can indicate which modifications are necessary to the transfer function. A rough estimate of

$$V_k - V_{ok} \quad \text{is} \quad r_{\alpha\hat{a}_o}(k)\,\frac{Sa_o}{S_\alpha}$$

Assume that the transfer function model is correct and that the "true" parameter estimates are substituted. Then the residual series a_t is a white noise series with approximate variance of $1/n$. The distributional properties of the auto-correlations are affected if we substitute parameter estimates in the model and the significance of individual crosscorrelations can be underestimated.

Haugh showed that on the assumption that the transfer function model is adequate then

$$Q_1 = n \sum_{k=0}^{M_1} r^2_{\alpha\hat{a}}(k) \quad \text{is approximately distributed as}$$

$$\chi^2_{M_1-T} \qquad (3.26)$$

and

$$Q_2 = n \sum_{k=-M_2}^{M_2} r^2_{\alpha a}(k) \quad \text{is approximately distributed as}$$

$$\chi^2_{2M_2-T} \qquad (3.27)$$

where $T = r+s$. Q_1 and Q_2 are compared against the upper 95% confidence point of the appropriate χ^2 distribution. This tests the overall adequacy of the transfer function model.

To check whether an unidirectional causality model rather than a feedback model is applicable we compare

$$Q_3 = n \sum_{k=-M_3}^{-1} r^2_{\alpha\hat{a}}(k) \quad \text{against} \quad \chi^2_{M_3}(0,95).$$

If $Q_3 > \chi^2_{M_3}(0,95)$ then a feedback mechanism exists.

3.8 FORECASTING

The transfer function model estimated is

$$\beta_t = V(B)\alpha_t + \psi(B)a_t$$

To generate forecasts we can rewrite the model in terms of the original series X_t and Y_t and then compute forecasts or β_t can be forecasted and these values can be used as the residuals in the univariate model for Y_t to compute the necessary forecasts.

Let

$$\hat{\beta}_t(\ell) = \sum_{j=0}^{\infty} V^0_{\ell+j}\alpha_{t-j} + \sum_{j=0}^{\infty} \psi^0_{\ell+j} a_{t-j} \tag{3.28}$$

be the ℓ-step ahead forecast from origin t for $\beta_{t+\ell}$.

Then

$$\beta_{t+\ell} - \hat{\beta}_t(\ell) = \sum_{j=0}^{\ell-1}(V_j\alpha_{t+\ell-j} + \psi_j a_{t+\ell-j})$$

$$+ \sum_{i=0}^{\infty}[(V_{\ell+i}-V^0_{\ell+i})\alpha_{t-i} + (\psi_{\ell+i}-\psi^0_{\ell+i})a_{t-i}]$$

so that

$$E[\beta_{t+\ell}-\hat{\beta}_t(\ell)]^2 = \sum_{j=0}^{\ell-1}(V_j^2\sigma_\alpha^2+\psi_j^2\sigma_a^2)$$

$$+ \sum_{i=0}^{\infty}[(V_{\ell+i}-V^0_{\ell+i})^2\sigma_\alpha^2 + (\psi_{\ell+i}-\psi^0_{\ell+i})^2\sigma_a^2] \tag{3.29}$$

The $E[\beta_{t+\ell}-\hat{\beta}_t(\ell)]^2$ is the Mean Square forecast error. It is minimized when $V^0_{\ell+i} = V_{\ell+i}$ and $\psi^0_{\ell+i} = \psi_{\ell+i}$. Thus the

mean square forecast $\hat{\beta}_t(\ell)$ is the conditional expectation of $\beta_{t+\ell}$ at time t.

To generate forecasts it is more convenient to write the model in the finite form

$$\beta_t = \delta^{-1}(B)\omega(B)\alpha_t + \phi^{-1}(B)\theta(B)a_t$$

Multiplying through by $\delta(B)\phi(B)$ gives

$$\delta(B)\phi(B)\beta_t = \omega(B)\phi(B)\alpha_t + \delta(B)\theta(B)a_t$$

or

$$\delta^*(B)\beta_t = \omega^*(B)\alpha_t + \theta^*(B)a_t \qquad (3.30)$$

where

$$\delta^*(B) = \delta(B)\phi(B)$$

$$\omega^*(B) = \omega(B)\phi(B)$$

$$\theta^*(B) = \delta(B)\theta(B)$$

Then

$$\begin{aligned}\hat{\beta}_t(\ell) &= E_t[\beta_{t+\ell}] \\ &= \delta_1^* E_t[\beta_{t+\ell-1}] + \ldots + \delta_{p+r}^* E_t[\beta_{t+\ell-p-r}] \\ &+ \omega_0^* E_t[\alpha_{t+\ell-b}] - \ldots - \omega_{p+s}^* E_t[\alpha_{t+\ell-b-p-s}] \\ &+ E_t[a_{t+\ell}] - \theta_1^* E_t[a_{t+\ell-1}] - \ldots - \theta_{q+r}^* E_t[a_{t+\ell-q-r}]\end{aligned} \qquad (3.31)$$

where

$$\begin{aligned}E_t[\beta_{t+j}] &= \beta_{t+j} && \text{if } j \leq 0 \\ &= \hat{\beta}_t(j) && \text{if } j > 0\end{aligned}$$

$$\begin{aligned}E_t[\alpha_{t+j}] &= \alpha_{t+j} && \text{if } j \leq 0 \\ &= \hat{\alpha}_t(j) && \text{if } j > 0\end{aligned}$$

$$\begin{aligned}E_t[a_{t+j}] &= a_{t+j} && \text{if } j \leq 0 \\ &= 0 && \text{if } j > 0\end{aligned}$$

Further $a_t = \beta_t - \hat{\beta}_{t-1}(1)$

The variance of the ℓ-step ahead forecast error is

$$V(\ell) = E[\beta_{t+\ell} - \hat{\beta}_t(\ell)]^2$$

$$= \sigma_\alpha^2 \sum_{j=b}^{\ell-1} V_j^2 + \sigma_a^2 \sum_{j=o}^{\ell-1} \psi_j^2 \qquad (3.32)$$

CHAPTER 4

THE APPLICATION OF THE BOX-JENKINS TECHNIQUES TO A SIMULTANEOUS ECONOMETRIC MODEL

4.1 INTRODUCTION

A major problem area in econometric analysis is the identification of causal relationships between variables. Regression analysis does not help the researcher to identify the direction of causality or the existence of feedback mechanisms. It is possible by means of transfer function model building to establish the direction of causality if it exists. It is preferable to carry out a causality analysis on variables of an econometric model based on sound economic theory. One has to keep in mind that one needs sufficient data (at least 50 observations) to be able to use the necessary programs for the analysis. Data observations on most monetary sector variables are available* from March 1965 on a monthly basis. The author used an econometric model established by Hurwitz (1977) in the analysis as it was based on economic theory and data was available for all the variables from November 1970.

Before the relationship between variables can be analysed, it is necessary to build univariate models for the different variables so as to take care of the autocorrelation present within each particular variable. The prewhitened series

---

* Reserve Bank Quarterly Bulletin

(the residuals obtained from the univariate models) are then crosscorrelated with each other and from the crosscorrelation function the causal relationship is established giving us our transfer function models. To utilize the knowledge obtained from the transfer function models, the author used these relationships in a regression analysis to establish alternative equations to the behavioural equations of Hurwitz' model for the purpose of forecasting. An analysis is then made of the different forecasts obtained.

4.2 THE ECONOMETRIC MODEL

In his Ph.D. thesis (Chapters 2 and 3) Hurwitz explains the money supply process in South Africa and formulates the following model consisting of four behavioural equations and one identity.

$$MB_t = R_t + NDA_t \tag{4.1}$$

$$R_t = f_1(R_{t-1}, NCDLED_{t-1}) \tag{4.2}$$

$$NDA_t = f_2(R_t, GD_{t-1}, NCDLTB_t) \tag{4.3}$$

$$TBR_t = f_3(R_{t-1}, NCD_{t-1}, NCDLTB_{t-1}) \tag{4.4}$$

$$MS_t = f_4(MB_t, E_{t-1}, NCD_{t-1}) \tag{4.5}$$

where

E_t = Electricity Generated at time t

GD_t = Government Deficit at time t

MB_t = Money Base at time t

MS_t = Money Supply at time t

NCD_t = 90 day NCD Rate at time t

$NCDLED_t$ = Difference between 90 day NCD Rate and the Euro Dollar Rate at time t

$NCDLTB_t$ = Difference between 90 day NCD Rate and the Treasury Bill Rate at time t

NDA_t = Net Domestic Assets at time t

R_t = Gold and Foreign Reserves at time t

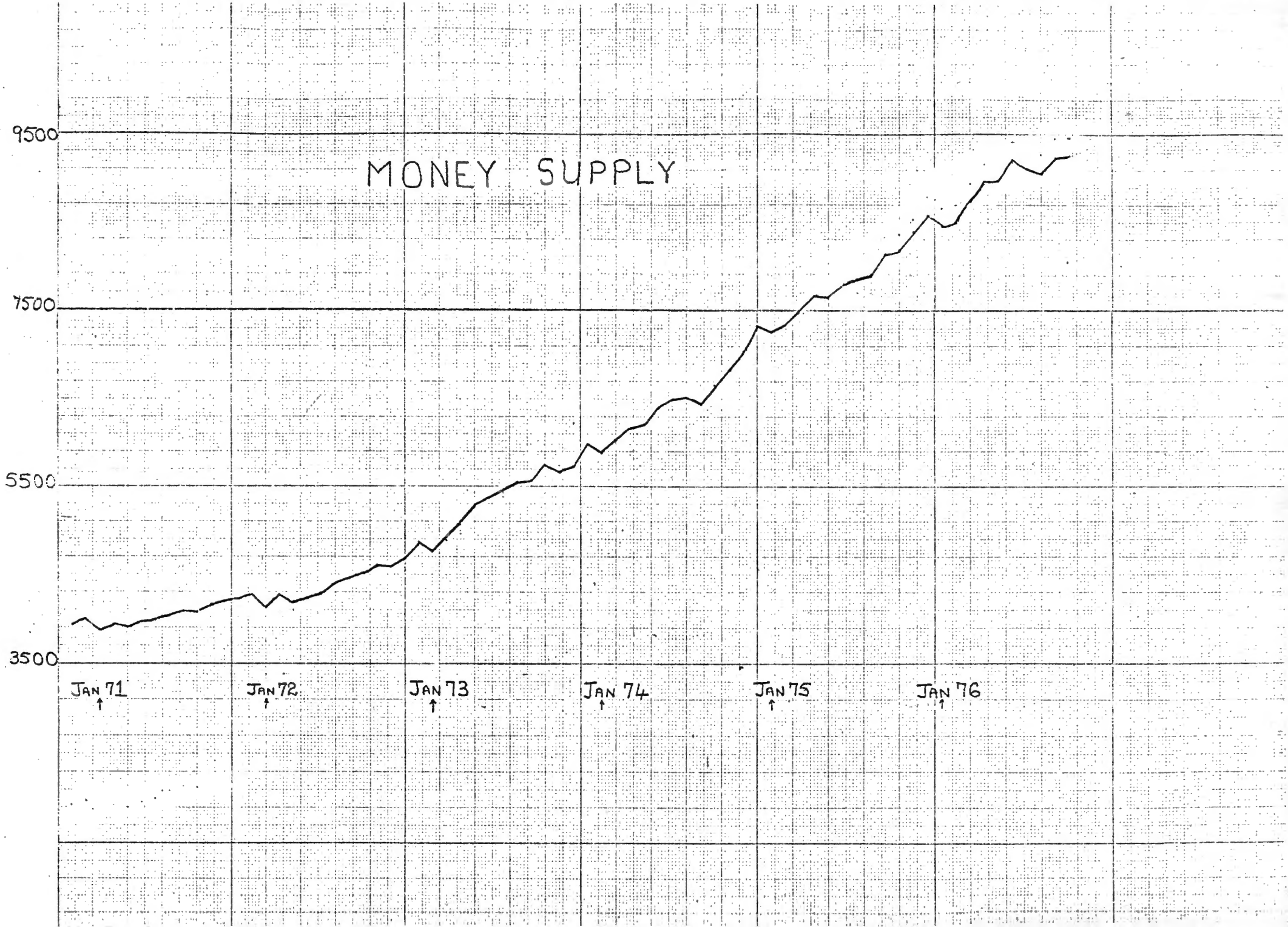
TBR_t = Treasury Bill Rate at time t

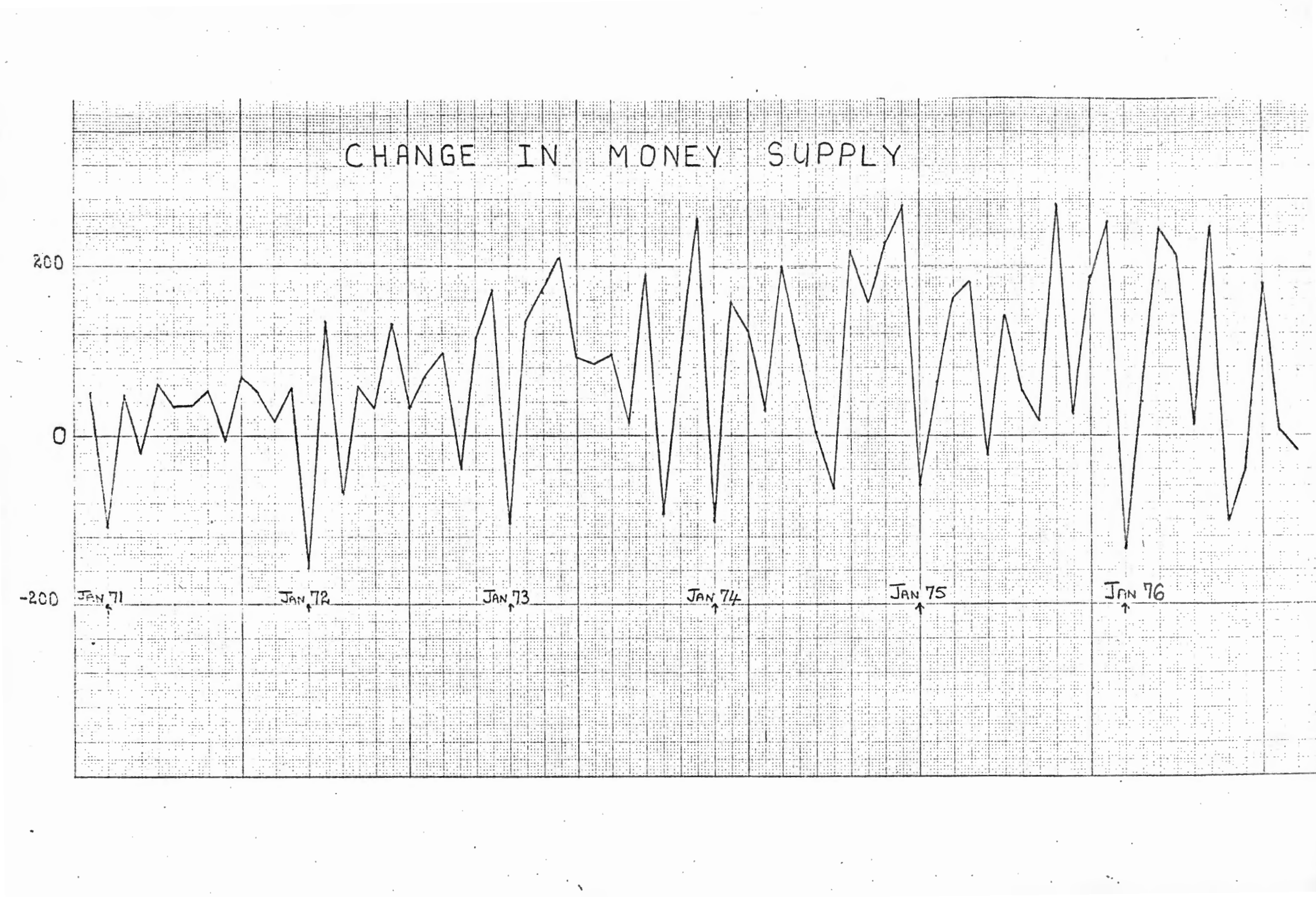
4.3 UNIVARIATE MODELS

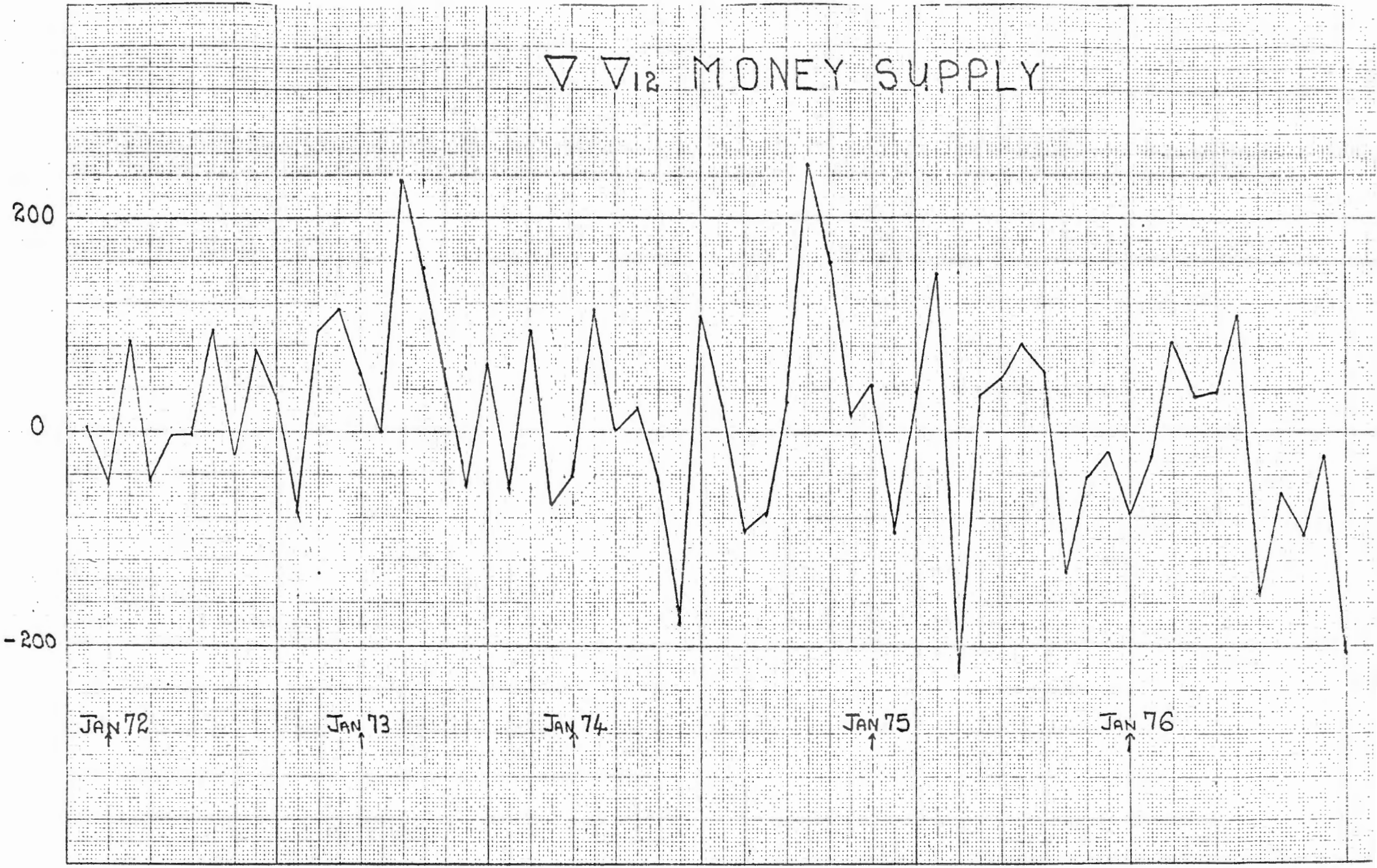
We build univariate models for the variables mentioned in Section 4.2. The procedure is explained in detail for Money Supply (MS).

4.3.1 GRAPHS

The plot of MS, ∇MS and $\nabla\nabla_{12}MS$ are given in Figures 4.1, 4.2 and 4.3 respectively. MS is nonstationary in level as it follows a linear trend pattern. To remove the nonstationarity it is necessary to difference MS once by order 1 giving ∇MS. ∇MS behaves seasonally as regular minima in January and maxima in September and December of each year occur. This pattern was not visible in MS as the linear trend damped the seasonal effect. Most values in ∇MS had the same sign as the value for the corresponding month in the previous year except 9 of which 4 occurred in August. The seasonal pattern behaved multiplicatively over the period 1971-1972 but became additive thereafter. To remove the seasonal pattern we difference ∇MS once by order 12 giving the series $\nabla\nabla_{12}MS$ which has neither a trend nor a seasonal pattern. Thus a stationary series is obtained.


MONEY SUPPLY
9500
7500
5500
3500
JAN 71
JAN 72
JAN 73
JAN 74
JAN 75
JAN 76


CHANGE IN MONEY SUPPLY
200
0
-200
JAN 71
JAN 72
JAN 73
JAN 74
JAN 75
JAN 76


∇ ∇12 MONEY SUPPLY
200
0
-200
JAN 72
JAN 73
JAN 74
JAN 75
JAN 76

4.3.2 TESTS FOR STATIONARITY

An indication of when the necessary level of differencing is reached is given by the Q-statistic, with $Q = n\sum_{k=1}^{p} r_k(a)^2$ distributed as χ_p^2 (see Chapter 2). A difference factor, say ∇, is necessary if it reduces the value of Q. The necessary level of differencing is reached when further differencing is increasing the value of Q. An overdifferenced series will usually include a Moving Average term in the identified model with a root close to the unit circle. Differencing has the same effect on the estimated variance of the series as it reduced the variance when necessary, but is increased by overdifferencing. The values obtained with $p = 24$ are

	Q-statistic	Variance
MS	579	3 158 084
∇MS	72	11 533
$\nabla\nabla_{12}$MS	28	8 497

Further $\chi_{24}^2(0,95) = 36,4$.

Thus $\nabla\nabla_{12}$MS does not differ significantly from a white noise series. The greatest reduction in Q and in the variance occurred by applying the difference factor ∇. Prothero suggested the following test for a seasonal time series. If ∇MS is a stationary series with variance σ_∇^2 then $\nabla\nabla_{12}$MS is also stationary with variance

$$\sigma_{\nabla\nabla_{12}}^2 = 2\,\sigma_\nabla^2(1-\rho_{12}) \qquad (4.6)$$

where ρ_{12} is the autocorrelation at lag 12 in the autocorrelation function of ∇MS.

Now $\sigma^2_{\nabla} = 11534,8$ $\sigma^2_{\nabla\nabla_{12}} = 8500,8$ $\rho_{12} = 0,51$

But $2\,\sigma^2_{\nabla}(1-\rho_{12}) = 11304,1 > 8500,8 = \sigma^2_{\nabla\nabla_{12}}$.

Thus we can conclude that ∇MS is nonstationary as the two values differ considerably.

4.3.3 AUTOCORRELATIONS AND PARTIAL AUTOCORRELATIONS

The autocorrelation and partial autocorrelation of $\nabla\nabla_{12}$MS are given in Table 4.1. The autocorrelation at lag 12 is significant while at lag 24 it is very small. Thus it cuts off after one lag of 12 indicating a moving average process. The partial autocorrelation function has a significant value at lag 12. Although the value at lag 24 is not significant it is much larger than the value in the autocorrelation function.

4.3.4 INITIAL ESTIMATE

An initial estimate for θ_{12} is obtained from

$$\rho_{12} = \frac{-\theta_{12}}{1+\theta^2_{12}} = 0,39 \qquad (4.7)$$

Thus

$$\theta^2_{12} - 2,564\ \theta_{12} + 1 = 0$$

Thus

$$\theta_{12} = 0,48 \quad \text{or} \quad 2,08.$$

The root 2,08 lies outside the invertible region and is ignored. Using $\theta_{12} = 0,48$ the partial autocorrelation for a moving average process of order 12 at lag 24 is

T A B L E 4.1

AUTOCORRELATION FUNCTION

DATA - MONEY SUPPLY

DIFFERENCE 1

MEAN OF THE SERIES = .13932+02
ST. DEV. OF SERIES = .92176+02
NUMBER OF OBSERVATIONS = 59

1- 6	.06	-.09	-.00	-.05	.13	-.07
ST.E.	.13	.13	.13	.13	.13	.13
7- 12	-.24	.07	.17	.19	-.05	-.39
ST.E.	.13	.14	.14	.15	.15	.15
13- 18	-.04	.06	-.08	-.11	-.07	.23
ST.E.	.17	.17	.17	.17	.17	.17
19- 24	.24	-.07	-.00	-.08	-.00	.07
ST.E.	.18	.18	.18	.18	.18	.18

MEAN DIVIDED BY ST. ERROR = .11610+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .28383+02 SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 24 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - MONEY SUPPLY

DIFFERENCE 1

MEAN OF THE SERIES = .13932+02
ST. DEV. OF SERIES = .92176+02
NUMBER OF OBSERVATIONS = 59

1- 6	.06	-.09	.01	-.05	.14	-.10
7- 12	-.21	.09	.15	.18	-.07	-.37
13- 18	-.08	.01	-.03	-.06	.04	.18
19- 24	.05	-.09	.15	.05	-.08	-.13

$$\Phi_{24\ 24} = \frac{-\theta_{12}^2(1-\theta_{12}^2)}{1-\theta_{12}^4} \qquad (4.8)$$

$$= -0,19$$

This value is larger than the value (-0,13) obtained in the partial autocorrelation function of $\nabla\nabla_{12}MS$ but of the same sign. This supports the conclusion of Section 4.3.3 that we have to fit a moving average process of order 12 to $\nabla\nabla_{12}MS$ as it is possible that the initial estimate of θ_{12} is too high.

4.3.5 ESTIMATION

The model obtained after estimation is

$$\nabla\nabla_{12}MS_t = a_t - 0,4005\ a_{t-12} \qquad (4.9)$$

The 95% confidence interval for θ_{12} is [0,126; 0,675]. Thus the estimate of θ_{12} differs significantly from zero and unity. The residual mean square equals 7465,2. If we look at the quantity $R^2 = 1 - \hat{\sigma}_a^2/\hat{\sigma}_{MS}^2$ where $\hat{\sigma}_a^2$ is an estimate of the residual variance and $\hat{\sigma}_{MS}^2$ an estimate of the variance of MS, then

$$R^2 = 1 - 6864/3\ 158\ 084$$

$$= 0,9978$$

so that the fitted model explains 99,78 per cent of the variation in MS. Most of the explanation is given by ∇MS as $R^2 = 1 - 11\ 533/3\ 158\ 084 = 0,9963$. The fitted model explains 40,48 per cent of the variation in ∇MS as $R^2 = 1 - 6864/11\ 533 = 0,4048$.

4.3.6 DIAGNOSTIC CHECKS

We now check that the fitted model is adequate. This requires that the residuals be uncorrelated, and normally distributed with mean zero. The individual residuals were examined and we find out of 59 residuals that 3 exceed twice the standard error and 18 exceed the standard error. This is very close to the spread of a normal distribution as the expected numbers are 2,65 and 18,73 respectively.

The most important check is the examining of the autocorrelation and partial autocorrelation functions which are given in Table 4.2. No individual value is significant. If the fitted model is appropriate, then $Q = 59\sum_{k=1}^{24} r_k^2(\hat{a})$ will be approximately distributed as χ^2_{23}. One degree of freedom is lost as we estimated θ_{12}.

Now $\chi^2_{23}(0,1) < Q = 15,68 < \chi^2_{23}(0,25)$.

Thus Q is not significant and we accept the model to be adequate.

4.3.7 OTHER UNIVARIATE MODELS

Univariate models for all the other variables were estimated and the results given in Table 4.3. The 95% confidence intervals for the parameters estimated in the models for E and TBR included zero or one. The models were used as the diagnostic checks pointed out that the inclusion of a moving average factor of order 1 in the model for E and an autoregressive and moving average factor of order 1 for TBR was

T A B L E 4.2

AUTOCORRELATION FUNCTION

DATA - THE ESTIMATED RESIDUALS - MODEL 1

ORIGINAL SERIES
MEAN OF THE SERIES = .24316+02
ST. DEV. OF SERIES = .82848+02
NUMBER OF OBSERVATIONS = 59

1- 6	.02	-.06	.05	-.10	.04	-.01
ST.E.	.13	.13	.13	.13	.13	.13
7- 12	-.18	.02	.20	.14	-.08	-.14
ST.E.	.13	.14	.14	.14	.14	.14
13- 18	-.06	-.04	-.04	-.10	-.04	.24
ST.E.	.15	.15	.15	.15	.15	.15
19- 24	.18	-.10	.03	-.06	-.04	.01
ST.E.	.16	.16	.16	.16	.16	.16

MEAN DIVIDED BY ST. ERROR = .22544+01

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .15672+02 SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 23 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - THE ESTIMATED RESIDUALS - MODEL 1

ORIGINAL SERIES
MEAN OF THE SERIES = .24316+02
ST. DEV. OF SERIES = .82848+02
NUMBER OF OBSERVATIONS = 59

1- 6	.02	-.06	.05	-.11	.05	-.03
7- 12	-.17	.01	.20	.15	-.11	-.15
13- 18	-.05	-.05	-.05	-.05	.00	.19
19- 24	.10	-.14	.06	.01	-.04	-.06

T A B L E 4.3

Variable	Estimated Model	95% C.I. of the parameters estimates	Estimated Residuals MSE	R^2	Q	d.f. p	$\chi^2_p(0,95)$
E	$E_t = E_{t-1}+a_t-0,1877\ a_{t-1}$	$-0,05<\theta_1<0,42$	4,59	0,9862	17,04	23	35,2
GD	$GD_t = 0,303GD_{t-1}+153,23+a_t$	$0,07<\phi_1<0,53$ $92,45<\theta_0<214,01$	31807,0	0,0881	17,55	22	33,9
MB	$\nabla MB_t = 0,5931\ \nabla MB_{t-12}+a_t$	$0,32<\phi_{12}<0,86$	1231,0	0,9823	8,33	23	35,2
MS	$\nabla\nabla_{12}MS_t = (1-0,4005B^{12})a_t$	$0,13<\theta_{12}<0,68$	7465,2	0,9978	15,68	23	35,2
NCD	$NCD_t = NCD_{t-1}+a_t$		1,34	0,8074	21,84	24	36,2
	$NCD_t = 0,905NCD_{t-1}+8,693+a_t$	$0,80<\phi_1<1,01$ $6,11<\theta_0<11,84$	1,31	0,8166	20,21	22	33,9
NCDLED	$\nabla NCDLED_t=0,3585\ \nabla NCDLED_{t-12}+a_t$	$0,12<\phi_{12}<0,60$	1,54	0,8323	13,93	23	35,2
NCDLTB	$NCDLTB_t = NCDLTB_{t-1}+a_t$		1,51	0,6304	21,09	24	36,4
	$(1-0,8194B)NCDLTB_t = 3,1549+a_t$	$0,68<\phi_1<0,96$ $1,59<\theta_0<4,72$	1,42	0,6634	16,88	22	33,9
NDA	$NDA_t = NDA_{t-1} + a_t$		7413,6	0,9101	12,84	24	36,4
R	$R_t = R_{t-1} + a_t$		5431,6	0,8606	17,10	24	36,4
TBR	$(1-0,6925B)\nabla TBR_t=(1-0,4526B)a_t$	$0,18<\phi_1<1,20$ $-0,17<\theta_1<1,08$	0,09	0,9493	15,20	22	33,9

necessary. A smaller mean square error was obtained with the above mentioned models.

A high degree of explanation was obtained in most models except for GD and NCDLTB. These two series were stationary and most of the variation was due to random fluctuation. All the models were adequate as no Q-statistic of goodness-of-fit was significant.

4.4 TRANSFER FUNCTION MODELS

The interrelationship between two series can be described by the cross correlation function provided that they are jointly covariance stationary. Covariance stationarity requires that the covariance between two series for all lags must be independent of time. The estimates of the cross correlations at all lags can be difficult to interpret if each of the series involved is autocorrelated to itself as the autocorrelation can inflate the variance of the crosscorrelation estimates above that expected when crosscorrelating white noise. Therefore it is preferable to prewhiten both series before cross correlating them.

4.4.1 RELATIONSHIP BETWEEN R AND NCDLED

The univariate models for R and NCDLED are listed in Table 4.3. Let PW(R) and PW(NCDLED) be the prewhitened series of R and NCDLED respectively. We crosscorrelate the prewhitened series and analyse the crosscorrelation

function which is given in Table 4.4.

4.4.1.1 CROSSCORRELATION FUNCTION

The analysis consists of comparing each lag of the crosscorrelation function with twice its approximate standard deviation $(1/\sqrt{n-k})$ where n is the number of observations ($n = 59$ in this case) and k the lag (in months). The value at lag 12 is significant which means that PW(NCDLED) leads PW(R) by 12 months. From lag 12 the crosscorrelations followed a second order decay pattern. A plot of the impulse response (which is proportional to the crosscorrelation function) and the step response function is given in Figure 4.4. From these plots the pattern is much more easily identified. As the crosscorrelations at lags 13, 14 etc. were not significant, the author decided to fit the following two models:

$$PW(R)_t = \omega_0 PW(NCDLED)_{t-12} + N_t \qquad (4.10)$$

and

$$PW(R)_t = \delta_1 PW(R)_{t-1} + \delta_2 PW(R)_{t-2} + \omega_0 PW(NCDLED)_{t-12} + N_t \qquad (4.11)$$

where ω_0 is an input lag parameter and δ_1 and δ_2 are output lag parameters. N_t is the generated noise. In the following sections models A and B are equations (4.10) and (4.11) respectively.

4.4.1.2 INITIAL ESTIMATES

The initial estimate of ω_0 in model A is v_{12}, the

T A B L E 4.4

CROSS CORRELATIONS

SERIES 1 - PREWHITENED RESIDUALS NCOLED
SERIES 2 - PREWHITENED GOLD AND FORFIGN RESERVES

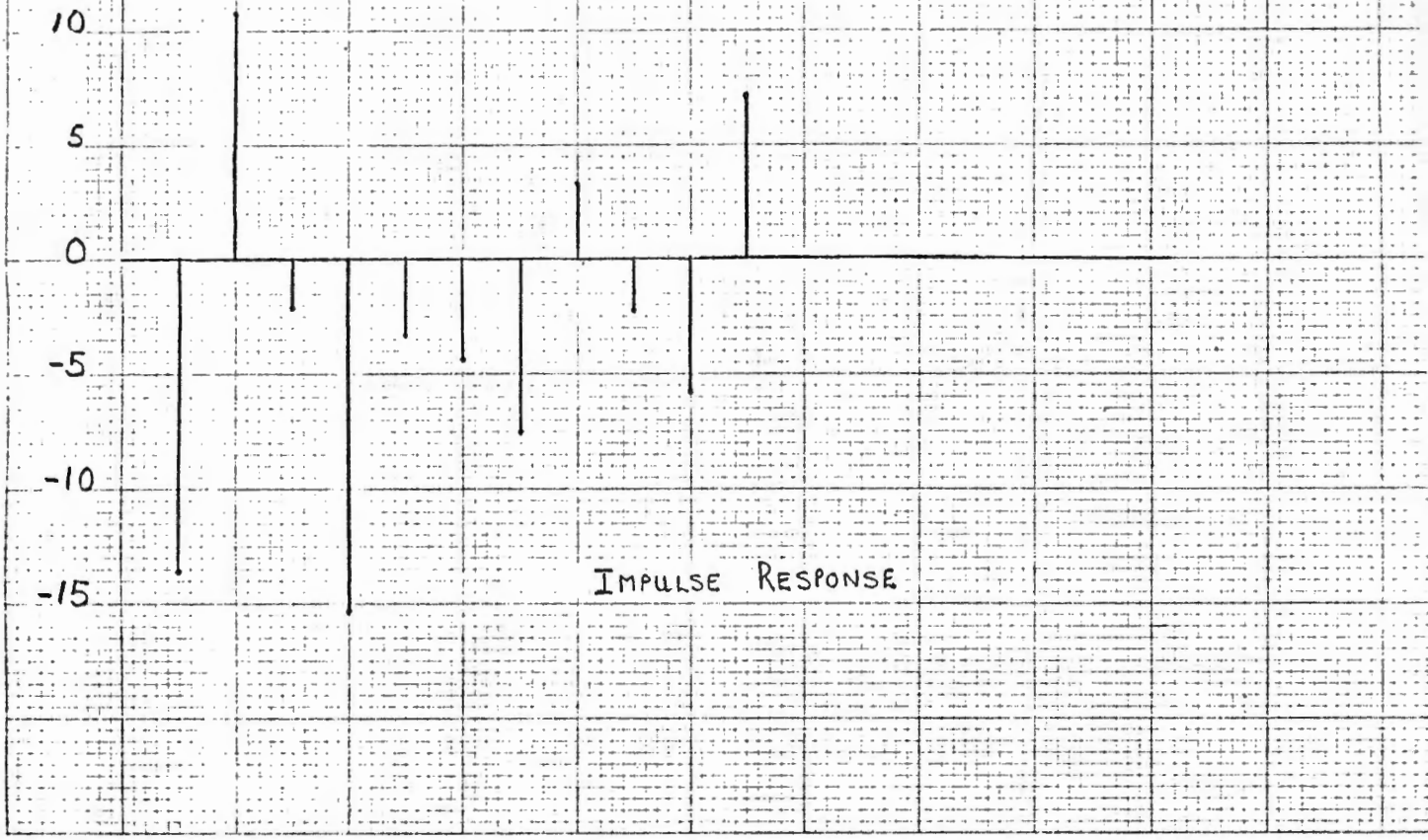
MEAN OF SERIES 1 = -.86610-01
ST. DEV. OF SERIES 1 = .12401+01
MEAN OF SERIES 2 = .48814+01
ST. DEV. OF SERIES 2 = .79285+02

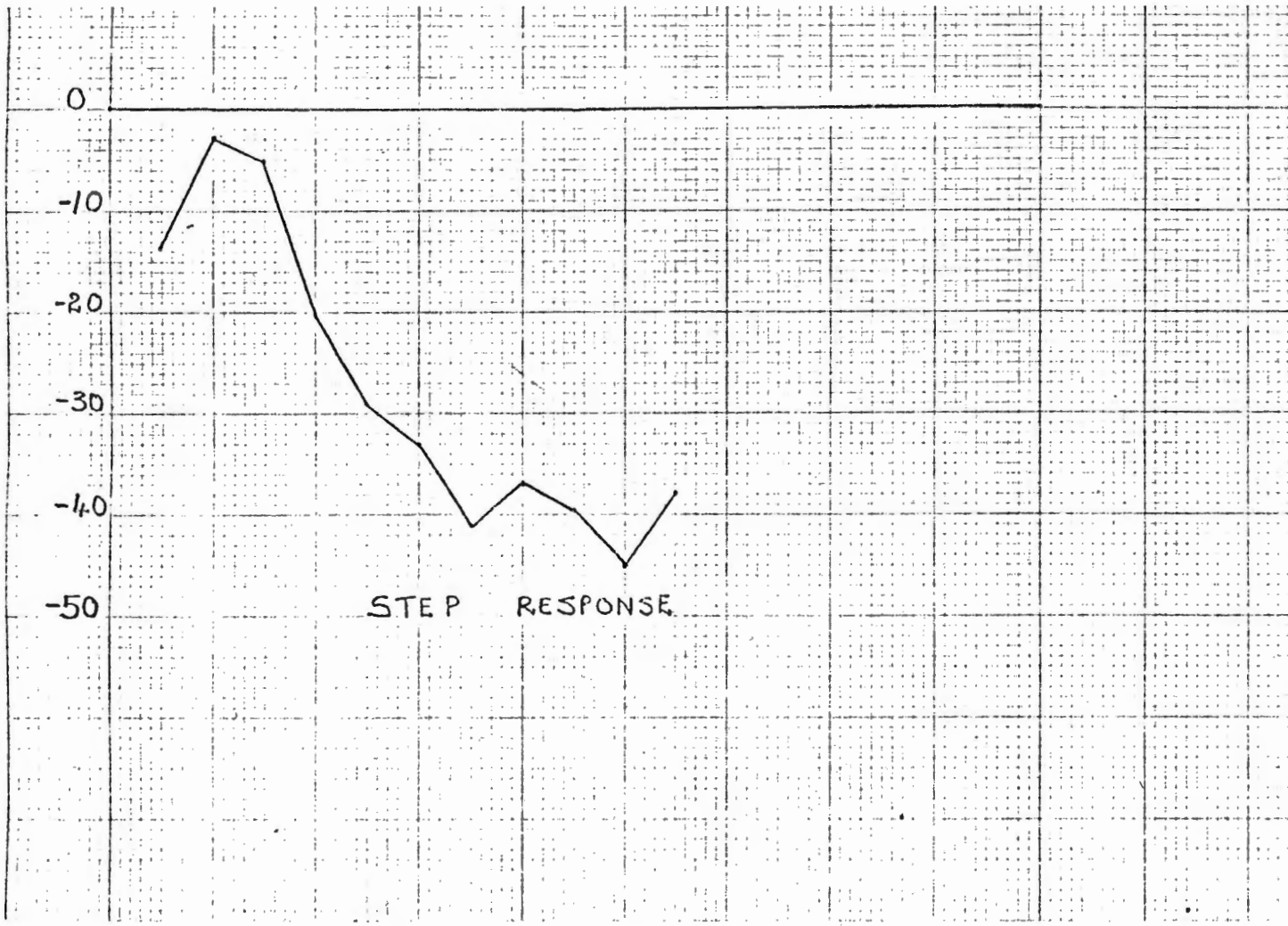
NUMBER OF LAGS ON SERIES 1	CROSS CORRELATION	NUMBER OF LAGS ON SERIES 2	CROSS CORRELATION
0	-.212	0	-.212
1	.165	1	-.245
2	-.032	2	-.202
3	-.240	3	.053
4	-.052	4	.034
5	-.069	5	.080
6	-.117	6	.090
7	.053	7	-.020
8	-.035	8	-.164
9	-.090	9	-.170
10	.112	10	-.075
11	-.123	11	.074
12	.292	12	-.046
13	.178	13	-.030
14	.129	14	.053
15	-.050	15	.093
16	-.161	16	-.006
17	-.077	17	.095
18	.278	18	.037
19	.145	19	.069
20	.129	20	.077

ESTIMATED IMPULSE RESPONSE WEIGHTS V(K)

K	V(K)
0	-13.526
1	10.578
2	-2.040
3	-15.372
4	-3.324
5	-4.435
6	-7.471
7	3.390
8	-2.262
9	-5.749
10	7.137
11	-7.875
12	18.646
13	11.369
14	8.278
15	-3.170
16	-11.543
17	-4.908
18	17.745
19	9.253
20	8.244

FIGURE 4.4





impulse response weight at lag 12, with estimated value 18,646. Initial estimates for ω_0, δ_1 and δ_2 in model B are obtained by solving the following set of equations.

$$\begin{aligned} v_{12} &= \omega_0 \\ v_{13} &= \delta_1 v_{12} + \delta_2 v_{11} \\ v_{14} &= \delta_1 v_{13} + \delta_2 v_{12} \\ v_{15} &= \delta_1 v_{14} + \delta_2 v_{13} \end{aligned} \qquad (4.12)$$

According to the theoretical function the estimate of v_{11} should be zero. The estimate (-7,875) did not differ significantly from zero. The initial estimates are

$$\hat{\omega}_0 = 18,646$$
$$\hat{\delta}_1 = 0,64$$
$$\hat{\delta}_2 = 0,06$$

which will be used to start the iteration routine in the estimation procedure.

4.4.1.3 GENERATED NOISE

To identify a model for the generated noise, N_t, we used the autocorrelation and partial autocorrelation function of the generated noise series (see Table 4.5). The approach is exactly the same as for univariate model building. No individual autocorrelations or partial autocorrelations were significant. Further $Q = 10,44 < 31,4 = \chi^2_{10}(0,95)$ so that we can accept it as a white noise series.

T A B L E 4.5

AUTOCORRELATION FUNCTION

DATA - THE GENERATED NOISE SERIES

ORIGINAL SERIES
MEAN OF THE SERIES = .44002+01
ST. DEV. OF SERIES = .76611+02
NUMBER OF OBSERVATIONS = 59

1- 5	.07	.17	-.03	.00	-.04
ST.E.	.13	.13	.13	.13	.13
6- 10	.15	-.05	.05	-.20	.05
ST.E.	.13	.14	.14	.14	.14
11- 15	-.12	.11	-.04	.03	-.13
ST.E.	.14	.15	.15	.15	.15
16- 20	.02	-.14	.02	-.05	-.02
ST.E.	.15	.15	.15	.15	.15

MEAN DIVIDED BY ST. ERROR = .44116+00

TO TEST WHETHER THIS SERIES IS WHITE NOISE, THE VALUE .10439+02 SHOULD BE COMPARED WITH A CHI-SQUARE VARIABLE WITH 20 DEGREES OF FREEDOM

PARTIAL AUTOCORRELATIONS

DATA - THE GENERATED NOISE SERIES

ORIGINAL SERIES
MEAN OF THE SERIES = .44002+01
ST. DEV. OF SERIES = .76611+02
NUMBER OF OBSERVATIONS = 59

1- 5	.07	.16	-.05	-.02	-.03
6- 10	.16	-.07	.00	-.19	.09
11- 15	-.07	.08	-.02	-.02	-.07
16- 20	.01	-.08	-.03	.02	-.07

4.4.1.4 ESTIMATION

For model A the estimated model is

$$PW(R)_t = 22,466 \quad PW(NCDLED)_{t-12} + a_t \qquad (4.13)$$

with Mean Square Error = 6487,4 and R^2 = 0,8363.
The 95% confidence interval for ω_0 is [3,823; 41,109] and therefore ω_0 differs significantly from zero at the 5% level.

The estimated model for model B is

$$PW(R)_t = 1,022\ PW(R)_{t-1} - 0,8069\ PW(R)_{t-2} + 12,925\ PW(NCDLED)_{t-12} + a_t \qquad (4.14)$$

with Mean Square Error = 5904,7 and R^2 = 0,8576.
The 95% confidence intervals are

$$[0,776 < \delta_1 < 1,268]$$

$$[-1,129 < \delta_2 < -0,485]$$

$$[1,226 < \omega_0 < 24,623]$$

Thus all the parameters estimated differed significantly from zero at the 5% level.

4.4.1.5 DIAGNOSTIC CHECKS

Diagnostic checks are made to test the adequacy of the estimated model. We first check that the generated noise series is correctly specified using the autocorrelation and partial autocorrelation functions of the estimated residuals.

For both model A and B no individual autocorrelation or

partial autocorrelation was significant. An overall check for adequacy of the functional form is to compare

$Q = n \sum_{k=1}^{t} r_{\hat{a}\hat{a}}(k)^2$ with an upper percentage point of the χ^2_{t-p-q} distribution where p is the number of autoregressive and q the number of moving average parameters included in the generated noise model. Now for model A

$Q = 8,09 < 31,4 = \chi^2_{20}(0,95)$ and for model B

$Q = 13,14 < 31,4 = \chi^2_{20}(0,95)$. Therefore the generated noise is correctly specified in both models.

We check the adequacy of the transfer function specification by examining the cross correlation function between the estimated residuals and the prewhitened input series PW(NCDLED). A comparison of $Q = n \sum_{k=0}^{s} r_{x\hat{a}}(k)^2$ with an upper percentage point of the $\chi^2_{(s+1)-(\ell+1)}$ distribution provides an overall check for the adequacy of the model. The number of corsscorrelations is $s+1$ and the number of parameters in the generated noise model is $\ell+1$. For model A

$$Q = 17,57 < 31,4 = \chi^2_{20}(0,95)$$

and therefore the model is adequate. For model B

$Q = 11,11 < 28,9 = \chi^2_{18}(0,95)$. This model is therefore also adequate.

4.4.1.6 COMPARISON OF THE UNIVARIATE MODEL WITH THE TRANSFER FUNCTION MODELS FITTED

Using the Mean Square Error (MSE) as the criterion of "best fit", model B was superior to model A as it exhibited a

lower MSE. Comparing model B with the univariate model, the latter model was superior. As all the parameters estimated in model B were significant at the 5% level, we would expect a lower MSE with model B than with the univariate model. In the univariate model one observation was used initially to calculate the estimated value for the second observation. We calculated 71 residuals over the period December 1970 to October 1976. In model B we used 13 values to calculate an estimate for the fourteenth observation so that 59 residuals were calculated over the period December 1971 to October 1976. If we break the calculation of the MSE for the univariate model into two parts, then the MSE's are as follows:

	RSS	df	MSE
November 1970 to November 1971	13 354	12	1113
November 1971 to October 1976	372 290	59	6310
November 1970 to October 1976	385 644	71	5432

[RSS = Residual Sum of Squares ; df = degrees of freedom]

If the residuals for the univariate and transfer function models were calculated over exactly the same period (November 1971 to October 1976), then the MSE of the transfer function model (5904,7) is lower than the MSE of the univariate model (6310) indicating an improvement on the univariate model fitted. In most cases where univariate and transfer function models will be fitted, the actual data used will cover the same period so that the residuals estimated will usually differ in the period covered. It is recommended that the period over which

the MSE is calculated in the univariate model is adjusted to the period actually used in the transfer function model if the residual structure differs from one period to another.

4.4.2 OTHER RELATIONSHIPS

The relationship between the different dependent and independent variables were analysed. The models estimated are listed in Table 4.6.

The model estimated between PW(NDA) and PW(R) is $PW(NDA)_t = -1,0796\ PW(R)_t + a_t/(1-0,6821B^{12})$. Further, the prewhitened model for NDA is

$$NDA_t - NDA_{t-1} = a_t = PW(NDA)_t$$

and for R is

$$R_t - R_{t-1} = a_t = PW(R)_t .$$

Replacing PW(NDA) and PW(R) in the transfer function model by the corresponding univariate models, we have $NDA_t - NDA_{t-1} = -1,0796\ (R_t - R_{t-1}) + a_t/(1-0,6821B^{12})$.

Thus

$$\nabla NDA_t = -1,0796\ \nabla R_t + a_t/(1-0,6821B^{12}) \qquad (4.15)$$

We conclude that the change in NDA is contemporaneously correlated with the change in R. As the relationship is negative an increase in R will lead to a decrease in NDA.

T A B L E 4.6

Variables	Estimated Model	95% C.I. of Parameters Estimates	MSE	R^2	Est.Residuals Q	d.f.	Crosscorrelation function Q	d.f.
R;NCDLED	$[1-1,022B+0,8069B^2]PW(R)_t =$ $12,925PW(NCDLED)_{t-12}+a_t$	$0,776<\delta_1<1,268$ $-1,129<\delta_2<-0,485$ $1,226<\omega_0<24,653$	5904,7	0,8576	13,14	20	11,11	18
	$PW(R)_t=22,47PW(NCDLED)_{t-12}+a_t$	$3,82<\omega_0<41,11$	6487,8	0,8363	8,09	20	13,67	19
NDA;R	$PW(NDA)_t=-1,0796PW(R)_t + \frac{a_t}{(1-0,6821B^{12})}$	$-1,19<\omega_0<-0,97$ $0,40<\phi_{12}<0,96$	1358,4	0,9840	9,39	19	25,15	18
NDA;NCDLTB	$PW(NDA)_t=(12,144-14,118B)$ $PW(NCDLTB)_t+a_t$	$-4,24<\omega_0<14,12$ $-2,32<\omega_1<30,56$	7042,0	0,9169	11,29	20	13,22	18
NDA,GD	No model							
TBR;R	$PW(TBR)_t=[-0,001185-0,00085B]$ $PW(R)_{t-1}+a_t$	$-0,0021<\omega_0<-0,0003$ $-0,000057<\omega_1<0,0017$	0,07893	0,9558	12,41	20	5,68	18

TABLE 4.6 (cont.)

Variables	Estimated Model	95% C.I. of Parameters Estimates	MSE	R^2	Est.Residuals Q	Est.Residuals d.f.	Crosscorrelation function Q	Crosscorrelation function d.f.
TBR;NCDLTB	$PW(NCDLTB)_t=[-1,3162-1,3167B^3 +0,8322B^4-0,8346B^5-1,2168B^6 -0,8593B^7]PW(TBR)_{t-10}+a_t$	$-2,26<\omega_0<-3,74$ $0,39<\omega_3<2,24$ $-1,75<\omega_4<0,08$ $-1,18<\omega_5<1,79$ $0,27<\omega_6<2,16$ $-0,09<\omega_7<1,81$	1,2368	0,7302	19,16	20	12,03	20
TBR;NCD	$PW(NCD)_t=[1,257-1,489B^{10} -1,005B^{17}]PW(TBR)_t+a_t$	$0,38<\omega_0<2,13$ $0,63<\omega_{10}<2,35$ $0,11<\omega_{17}<1,90$	1,0867	0,8516	17,58	20	16,73	16
MS;NCD	$PW(MS)_t=[16,353+20,709B^{12}]$ $PW(NCD)_{t-3} + a_t$	$-0,74<\omega_0<33,45$ $-39,34<\omega_{12}<-2,08$	6099,4	0,9981	16,77	20	22,48	17
MS;E	$PW(MS)_t=11,76PW(E)_{t-4}+a_t$	$1,96<\omega_0<21,56$	5713,5	0,9982	21,55	20	13,42	16
MS;MB	No model							

The functional relationship between NDA and R,GD and NCDLTB is given in Section 4.2. If we write equation (4.3) in a difference form, then

$$\nabla NDA_t = f_5(\nabla R_t, \nabla GD_{t-1}, \nabla NCDLTB) \qquad (4.16)$$

The Ordinary Least Squares estimate of equation (4.3) is

$$NDA_t = 535{,}66 - 0{,}6504\ R_t + 0{,}5352\ GD_{t-1} + 60{,}004\ NCDLTB_t$$

(see Section 4.5 for further details).
If we rewrite this equation in terms of first differences we find that the sign for the coefficient of ∇R in the transfer function model is the same as in the regression model.

The model estimated for PW(NDA) and PW(NCDLTB) is

$$PW(NDA)_t = (12{,}144 - 14{,}118B)\ PW(NCDLTB)_t + a_t.$$

The univariate model for NCDLTB is

$$NCDLTB_t - NCDLTB_{t-1} = a_t = PW(NCDLTB)_t .$$

If $PW(NDA)_t$ and $PW(NCDLTB)_t$ are replaced by their univariate models, then

$$NDA_t - NDA_{t-1} = 12{,}144(NCDLTB_t - NCDLTB_{t-1}) - 14{,}118(NCDLTB_{t-1} - NCDLTB_{t-2}) + a_t .$$

Thus

$$\nabla NDA_t = 12{,}144\ \nabla NCDLTB_t - 14{,}118\ \nabla NCDLTB_{t-1} + a_t \qquad (4.17)$$

In both the transfer function model and the regression model the coefficient of $\nabla NCDLTB$ has the same sign.

No transfer function model could be estimated between NDA and GD as we find no significant crosscorrelations.

The relationship between PW(TBR) and PW(R) is

$$PW(TBR)_t = -0,001185\ PW(R)_{t-1} - 0,00085\ PW(R)_{t-2} + a_t.$$

The univariate model for TBR is

$$(1-0,6925B)\ \nabla TBR_t = a_t - 0,4526\ a_{t-1}.$$

In terms of the prewhitened series it is

$$a_t = (1-0,6925B)(1-0,4526B)^{-1}(1-B)TBR_t = PW(TBR)_t \qquad (4.18)$$

If we replace PW(TBR) by equation (4.18) and PW(R) by ∇R in the transfer function model, then

$$(1-0,6925B)(1-0,4526B)^{-1}(1-B)TBR_t = -0,001185\ \nabla R_{t-1} - 0,00085\ \nabla R_{t-2} + a_t$$

so that

$$\nabla TBR_t = -0,001185\ \nabla R_{t-1} - 0,001135\ \nabla R_{t-2} - 0,000401\ \nabla R_{t-3} + (1-0,4526B)(1-0,6925B)^{-1}a_t \qquad (4.19)$$

The regression equation estimated for TBR (see Section 4.5) is

$$TBR_t = 0,7469 - 0,0006413\ R_{t-1} + 0,9209\ NCD_{t-1} - 0,8596\ NCDLTB_{t-1}$$

If we rewrite the equation in terms of differences, we find that the sign of the coefficient of ∇R_{t-1} agrees with the sign of the coefficient in the transfer function model.

The transfer function model identified between PW(TBR) and PW(NCDLTB) gave PW(TBR) as a leading series for PW(NCDLTB). This contradicts the regression relationship as NCDLTB leads TBR by one lag. No economic explanation can be given why a significant crosscorrelation occurred at lag 10. A similar result was obtained when crosscorrelating PW(TBR) with PW(NCD). There is also no economic justification for lags 10 and 17 in this model.

The prewhitened MS and MB series were independent. This differs from the regression equation in which they were contemporaneously correlated. The regression equation estimated for MS (see Section 4.5) is

$$MS_t = -4375{,}6 + 3{,}7636\ MB_t + 50{,}2351\ E_{t-1} - 17{,}57\ NCD_{t-1}.$$

In the transfer function model the series PW(MS) lagged PW(E) by 4 periods. The estimated model is

$$PW(MS)_t = 11{,}76\ PW(E)_{t-4} + a_t$$

where

$$PW(MS)_t = a_t = (1-0{,}4005B^{12})^{-1}\ \nabla\nabla_{12}MS_t \qquad (4.20)$$

and

$$PW(E)_t = a_t = (1-0{,}1877B)^{-1}\ \nabla E_t \qquad (4.21)$$

If PW(MS) and PW(E) are replaced in the transfer function model by equations (4.19) and (4.20) respectively then the model becomes

$$(1-0{,}4005B^{12})^{-1}\ \nabla\nabla_{12}MS_t = 11{,}76\ (1-0{,}1877B)^{-1}\ \nabla E_{t-4} + a_t\ .$$

Rewriting it in terms of ∇MS_t only, gives us

$$\nabla MS_t = 11,76\ (1-0,1877B)^{-1}(1-0,4005B^{12})(1-B^{12})^{-1}$$
$$\nabla E_{t-4} + (1-0,4005B^{12})(1-B^{12})^{-1}\ a_t$$
$$= 11,76\ [1+0,1877B+0,0352B^2+0,5995B^{12}+...]\ \nabla E_{t-4}$$
$$+ (1-0,4005B^{12})(1-B^{12})^{-1}\ a_t \qquad (4.22)$$

The coefficients of all lags in ∇E are positive. The lags differ from lag 1 used in the regression equation but do have the same sign.

The cross correlation analysis between PW(MS) and PW(NCD) showed that the latter series led the former by 3 periods. If we rewrite the model estimated

$$PW(MS)_t = [16,353 + 20,709B^{12}]\ PW(NCD)_{t-3} + a_t$$

in terms of ∇MS and ∇NCD we obtain the following model

$$\nabla MS_t = 16,353\ \nabla NCD_{t-3} + 30,51\ \nabla NCD_{t-15} + 22,22\ \nabla NCD_{t-27}$$
$$+ ... + (1-0,4005B^{12})(1-B^{12})^{-1}\ a_t \qquad (4.23)$$

In this case both the lags and the sign of the coefficients of the parameters estimated differed. The positive sign of the coefficients can be explained but not the different lag periods. If the interest rate increases, the banks demand for reserves will decrease, banks will expand credit facilities, the Brunner-Meltzer multiplier will increase and thus money supply will increase.

4.5 REGRESSION EQUATIONS

The regression equations were estimated using Ordinary Least Squares. The results are represented below with the relevant ordinary least square summary statistics which are

(i) t-statistics, which are given below each coefficient. Each t-statistic has n-k degrees of freedom where k is the number of explanatory variables plus one, and n the number of observations.

(ii) R^2 - the multiple coefficient of determination.

(iii) DW - the Durbin-Watson statistic.

(iv) SE - the standard error of the estimate.

(v) df - degrees of freedom.

$$MS_t = \underset{[-15,05]}{-4375,6} + \underset{[13,83]}{3,7636}\ MB_t + \underset{[12,04]}{50,2351}\ E_{t-1} - \underset{[-1,57]}{17,57}\ NCD_{t-1} \qquad (4,24)$$

(R^2 = 0,9877 SE = 200,63 DW = 0,66 df = 67)

$$NDA_t = \underset{[4,47]}{535,66} - \underset{[-5,13]}{0,6504}\ R_t + \underset{[3,84]}{0,5352}\ GD_{t-1} + \underset{[4,75]}{60,004}\ NCDLTB_t \qquad (4.25)$$

(R^2 = 0,4863 SE = 211,31 DW = 0,44 df = 67)

$$R_t = \underset{[1,31]}{53,70} + \underset{[19,37]}{0,9323}\ R_{t-1} + \underset{[0,003]}{0,008637}\ NCDLED_{t-1} \qquad (4,26)$$

(R^2 = 0,8670 SE = 74,04 DW = 1,76 df = 68)

$$TBR_t = \underset{[1,88]}{0,7469} - \underset{[-2,23]}{0,0006413}\ R_{t-1} + \underset{[19,70]}{0,9209}\ NCD_{t-1} - \underset{[-14,47]}{0,8596}\ NCDLTB_{t-1} \qquad (4,27)$$

(R^2 = 0,8926 SE = 0,4456 DW = 1,76 df = 67)

The one-sided 95% point of the t-statistic with 67 degrees of freedom is 1,6695 (or -1,6695 for a left sided test). The constant in equation (4.26) was non-significant. Non-significant coefficients were obtained for NCD_{t-1} in the money supply equation (4.24) and $NCDLED_{t-1}$ in the gold and foreign reserves equation (4.26). In the money supply equation (4.24) the non-significance of the coefficient of NCD_{t-1} could be caused by multicollinearity as MB_t and E_{t-1} were highly correlated. The correlation coefficient matrix is

	MB	E_{-1}	NCD_{-1}	MS
MB	1,00			
E_{-1}	0,94	1,00		
NCD_{-1}	0,60	0,56	1,00	
MS	0,98	0,98	0,57	1,00

The econometric model of Section 4.2 was structured to explain the economic behaviour over the period 1970 to 1974. As this analysis covered the period until October 1976 it is possible that the above mentioned explanatory variables did not influence the dependent variables as specified.

For the Durbin-Watson statistics the following 95% significance points for 70 observations apply:

k = 2		k = 3		k = 4	
d_L	d_u	d_L	d_u	d_L	d_u
1,55	1,67	1,52	1,70	1,49	1,74

where k is the number of explanatory variables

d_L is the lower significance bound

d_u is the upper significance bound

For equations (4.24) the Durbin-Watson statistic takes the value 0,66 which is less than 1,52, the lower significance bound with 3 explanatory variables for the statistic. Thus a significant positive correlation of the first order is present. For equation (4.25) we have that

$$DW = 0,44 < 1,52 = d_L$$

Thus a significant positive correlation of the first order is present.

The author did not correct for multicollinearity by means of Ridge-regression and for autocorrelation using the Cochran-Orcutt or the Scanning technique as it was not the intention of this study to discuss regression analysis in detail. For forecasting purposes we built univariate models for the residuals of MS and NDA and included a forecasted residual value in the forecast produced. The univariate model for the residuals of the regression on MS (4.24) is

$$(1 - 0,6352B)(1 + 0,2057B^8)\, X_t = a_t \qquad (4.28)$$

where

X_t is the residuals of the regression on MS and a_t the residuals obtained after fitting the above univariate model.

The 95% confidence intervals for the parameter estimates are

$$0,42 < \phi_1 < 0,84$$

$$-0,52 < \phi_8 < 0,10$$

Further MSE = 24108,0 and R^2 = 0,3940.

For the residuals of the regression on NDA (4.25) the univariate model is

$$(1 + 0,3267B^7 - 0,6093B^{23})(1-B)\ X_t = (1 - 0,4918B)a_t \qquad (4.29)$$

where X_t are the residuals of the regression on NDA and a_t the residuals obtained after fitting the above univariate model. The 95% confidence intervals for the parameter estimates are:

$$-0,647 < \phi_7 < -0,007$$

$$0,222 < \phi_{23} < 0,996$$

$$0,216 < \theta_1 < 0,768$$

Further MSE = 16878,0 and R^2 = 0,64.

4.6 SPECIFICATION OF EQUATIONS USING THE RESULTS OF TRANSFER FUNCTION MODEL BUILDING

In Section 4.4 we established relationships between the different explanatory and dependent variables. Utilizing this information the following equations are specified:

$$\nabla R_t = f_6(\nabla NCDLED_{t-12}) \qquad (4.30)$$

$$\nabla NDA_t = f_7(\nabla R_t, \nabla NCDLTB_t, \nabla NCDLTB_{t-1}) \qquad (4.31)$$

$$\nabla TBR_t = f_8(\nabla R_{t-1}, \nabla NCD_t) \qquad (4.32)$$

$$\nabla MS_t = f_9(\nabla MS_{t-12}, \nabla NCD_{t-3}, \nabla NCD_{t-15}, \nabla E_{t-4}) \qquad (4.33)$$

f_6, f_7, f_8, f_9 are linear functions of the explanatory variables.

The purpose of constructing the above equations is to compare their forecasting performance with those of the univariate models (Section 4.3), transfer function models (Section 4.4) and the regression models of Section 4.5. As explained in Section 4.2.2 it is not possible to interpret all the above results from an economic viewpoint. If the forecasting performance of the above mentioned equations and the transfer function models are not superior to the other two methods, then the established relationships should be treated as chance occurences. The equations (4.30) to (4.33) are estimated by Ordinary Least Sqaures. The results are:

$$\underset{}{\nabla R_t} = \underset{[0,46]}{4,5457} + \underset{[2,46]}{18,1667}\ \nabla NCDLED_{t-12} \qquad (4.34)$$

(R^2 = 0,0802 SE = 76,69 DW = 1,80 df = 57)

$$\nabla TBR_t = \underset{[0,90]}{0,035} - \underset{[-2,03]}{0,0010}\ \nabla R_{t-1} + \underset{[2,78]}{0,0925}\ \nabla NCD_t \qquad (4.35)$$

(R^2 = 0,1987 SE = 0,2991 DW = 1,81 df = 56)

$$\nabla MS_t = \underset{[2,95]}{41,8934} + \underset{[4,55]}{0,5268}\ \nabla MS_{t-12} + \underset{[1,45]}{13,6199}\ \nabla NCD_{t-3} \qquad (4.36)$$

(R^2 = 0,5052 SE = 78,15 DW = 2,15 df = 51)

$$\nabla NDA_t = \underset{[2,39]}{13,2127} - \underset{[-14,38]}{1,0023}\ \nabla R_t + \underset{[0,88]}{3,6877}\ \nabla NCDLTB_t - \underset{[-1,28]}{5,3719}\ \nabla NCDLTB_{t-1} \qquad (4.37)$$

(R^2 = 0,8049 SE = 41,27 DW = 2,58 df = 52)

The constants in equations (4.34) and (4.35) were non-significant. The coefficient of ∇NCD_{t-3} in equation (4.36) and the coefficients of $\nabla NCDLTB_t$ and $\nabla NCDLTB_{t-1}$ in equation (4.37) were non-significant. As no multicollinearity was present in any of these two equations, which could influence the value of the t-statistics, it means that the influence of the above mentioned explanatory variables on the dependent variables are too weak to be statistically significant and will be ignored. The corresponding correlation coefficient matrices are given below:

	∇MS_{-12}	∇NCD_{-3}	∇NCD_{-15}	∇E_{-4}	∇MS
∇MS_{-12}	1,00	0,25	0,34	-0,01	0,61
∇NCD_{-3}		1,00	0,17	0,35	0,40
∇NCD_{-15}			1,00	-0,03	0,48
∇E_{-4}				1,00	0,25
∇MS					1,00

	∇R	$\nabla NCDLTB$	$\nabla NCDLTB_{-1}$	∇NDA
∇R	1,00	-0,16	0,16	-0,90
$\nabla NCDLTB$		1,00	-0,13	0,20
$\nabla NCDLTB_{-1}$			1,00	-0,23
∇NDA				1,00

We re-estimate equations (4.36) and (4.37) ignoring ∇NCD_{t-3}, $\nabla NCDLTB_t$ and $\nabla NCDLTB_{t-1}$ as explanatory variables. The results were:

$$\nabla MS_t = \underset{[2,69]}{37,788} + \underset{[4,96]}{0,5652}\ \nabla MS_{t-12} + \underset{[3,09]}{30,4635}\ \nabla NCD_{t-15} + \underset{[2,75]}{13,80}\ \nabla E_{t-4} \qquad (4.38)$$

($R^2 = 0,4949$ SE = 78,96 DW = 2,13 df = 52)

$$\nabla NDA_t = \underset{[2,39]}{13,2649} - \underset{[-14,98]}{1,0264}\ \nabla R_t \qquad (4.39)$$

($R^2 = 0,8025$ SE = 41,52 DW = 2,58 df = 54)

The equations gave a very low explanation of the variation in the variables ∇TBR and ∇R but satisfactory explanation in ∇MS and ∇NDA. These results should be looked at in the context that all these series were nonstationary and that the first differences explained a high percentage of the variation in the original series. The percentage explanations given by the first differences were

	R^2
MS	0,9963
TBR	0,9451
NDA	0,9101
R	0,8606

(these figures were based on 71 observations).

The Durbin-Watson statistic for ∇NDA (4.37) showed significant negative autocorrelation (DW = 2,58). The transfer function model (Table 4.6) included autocorrelation of order 12 which will be difficult, if not impossible, to remove by the available techniques which only correct first and second order autocorrelation.

4.7 FORECASTING PERFORMANCE

One step ahead forecasts were produced for the dependent variables NDA, MS, TBR and R over the period November 1976 to March 1978. The forecasts are listed in Tables 4.7 to 4.10. In the regression equations and transfer function models the forecasts of independent variables were obtained from the corresponding univariate model.

4.7.1 COMPARISONS OF FORECASTING METHODS

To evaluate the forecasts produced by the various methods we analyse the errors that occurred. Various criteria could be used to compare the errors, but it is generally considered that the absolute error and quadratic error functions are preferable.

T A B L E 4.7

MONEY SUPPLY

Origin	Actual Data	1-step ahead forecasts				
			Transfer Functions*		Regressions	
		Univariate	E	NCD	Eq(4.24)	Eq(4.38)
Nov 76	9368	9492	9467	9470	8971	9317
Dec 76	9115	9255	9296	9254	9175	9367
Jan 77	9174	9177	9185	9199	9417	9257
Feb 77	9146	9376	9352	9373	8904	9304
Mar 77	9427	9330	9353	9322	8709	9234
Apr 77	9548	9452	9580	9410	8933	9712
May 77	9799	9745	9747	9726	9560	9526
June 77	9677	9759	9693	9712	9124	9665
July 77	9613	9653	9571	9646	9551	9741
Aug 77	9853	9818	9877	9816	9084	9745
Sept 77	9703	9874	9893	9840	9901	9841
Oct 77	9848	9764	9812	9758	9875	9877
Nov 77	10012	10024	10030	10065	9727	9748
Dec 77	9726	9815	9824	9843	9957	10052
Jan 78	9719	9786	9941	9798	9414	9910
Feb 78	9873	9783	9786	9782	9527	9896
C_1		88,25	88,19	87,44	330,63	149,56
C_2		10971,63	12814,31	10665,56	157372,88	31105,69

* The symbols below indicate the leading series in the transfer function model

T A B L E 4.8

NET DOMESTIC ASSETS

Origin	Actual Data	1-step ahead Forecasts				
		Univariate	Transfer Functions* R	NCDLTD	Regressions Eq(4.25)	Eq(4.39)
Nov 76	766	792	826	795	724	767
Dec 76	803	766	809	758	873	785
Jan 77	880	803	762	807	750	832
Feb 77	835	880	844	875	926	969
Mar 77	751	835	833	843	477	844
Apr 77	762	751	647	754	954	653
May 77	743	762	721	767	982	790
June 77	786	743	801	745	773	750
July 77	776	786	846	789	713	793
Aug 77	870	776	818	777	844	833
Sept 77	836	870	862	865	792	895
Oct 77	925	836	822	838	1100	831
Nov 77	958	925	917	928	**	961
Dec 77	905	958	970	957		947
Jan 78	899	905	875	901		905
Feb 78	1026	899	893	890		920
C_1		49,13	58,81	49,63	113,25	53,13
C_2		3588,88	5084,94	3744,50	19856,75	4477,50

*The symbols below indicate the leading series used in the transfer function model.

**The data for government deficit is not available.

T A B L E 4.9

GOLD AND FOREIGN RESERVES

Origin	Actual Data	1-step ahead Forecasts				
			Transfer Function		Regressions	
		Univariate	Eq(4.13)	Eq(4.14)	Eq(4.26)	Eq(4.34)
Nov 76	734	740	775	744	770	733
Dec 76	719	734	755	738	769	724
Jan 77	645	719	777	724	773	747
Feb 77	649	645	655	655	666	688
Mar 77	757	649	625	659	621	656
Apr 77	743	757	730	759	753	715
May 77	749	743	761	746	752	704
June 77	755	749	740	752	738	739
July 77	712	755	763	758	775	780
Aug 77	700	712	681	717	696	728
Sept 77	682	700	684	706	690	687
Oct 77	660	682	683	690	679	656
Nov 77	636	660	668	669	639	649
Dec 77	623	636	614	647	634	633
Jan 78	631	623	611	635	636	622
Feb 78	634	631	611	642	625	621
C_1		23,50	35,38	30,44	26,25	32,44
C_2		1331,50	2745,00	1949,06	1390,13	2750,56

T A B L E 4.10

TREASURY BILL RATE

Origin	Actual Data	1-step ahead Forecasts			
		Univariate	Transfer Functions* R	Regressions Eq(4.27)	Eq(4,35)
Nov 76	7,74	7,72	7,69	7,49	7,69
Dec 76	7,82	7,75	7,75	7,55	7,83
Jan 77	7,87	7,84	7,88	7,61	7,85
Feb 77	7,88	7,89	8,01	7,77	8,02
Mar 77	7,86	7,89	7,97	7,70	7,86
Apr 77	7,82	7,86	7,75	7,60	7,76
May 77	7,85	7,81	7,76	7,55	7,83
June 77	7,88	7,85	7,88	7,56	7,87
July 77	7,88	7,89	7,90	7,57	7,89
Aug 77	7,86	7,88	7,95	7,59	7,95
Sept 77	7,92	7,86	7,93	7,61	7,94
Oct 77	7,92	7,93	7,98	7,66	7,93
Nov 77	7,90	7,93	7,99	7,67	7,96
Dec 77	7,93	7,90	7,97	7,67	7,90
Jan 78	7,95	7,94	8,00	7,72	7,98
Feb 78	7,98	7,96	7,98	7,77	8,04
C_1		0,0288	0,0494	0,2369	0,0388
C_2		0,00111	0,00462	0,06023	0,00278

* R is the leading series used for the transfer function model

With the absolute error function as our criterion, a set of forecasts generated from a model outperforms other models used if it has a lower value for C_1 where

$$C_1 = \frac{1}{N} \sum_{t=1}^{N} |e_t|$$

is the average absolute error over the forecasting period and N is the number of forecasts. The absolute error function applies a weight equal to the magnitude of an error, and it is therefore possible to specify which model forecasts the "closest" to the data. If a quadratic error function is used more weight is given to large than to small errors as the weight equals the error squared. Thus the "closest" forecast to the data will not necessarily minimize the function C_2 where

$$C_2 = \frac{1}{N} \sum_{t=1}^{N} e_t^2$$

is the average quadratic error.

The quadratic error function is suitable as a criterium as it corresponds to the least square criterium, is mathematically more tractable than other criteria and is a reasonable assumption. The values obtained for C_1 and C_2 are listed in Tables 4.7 to 4.10. We find that the univariate models for R, NDA and TBR and the transfer function model for MS with NCD as leading series were superior to the other methods using both the above mentioned criteria.

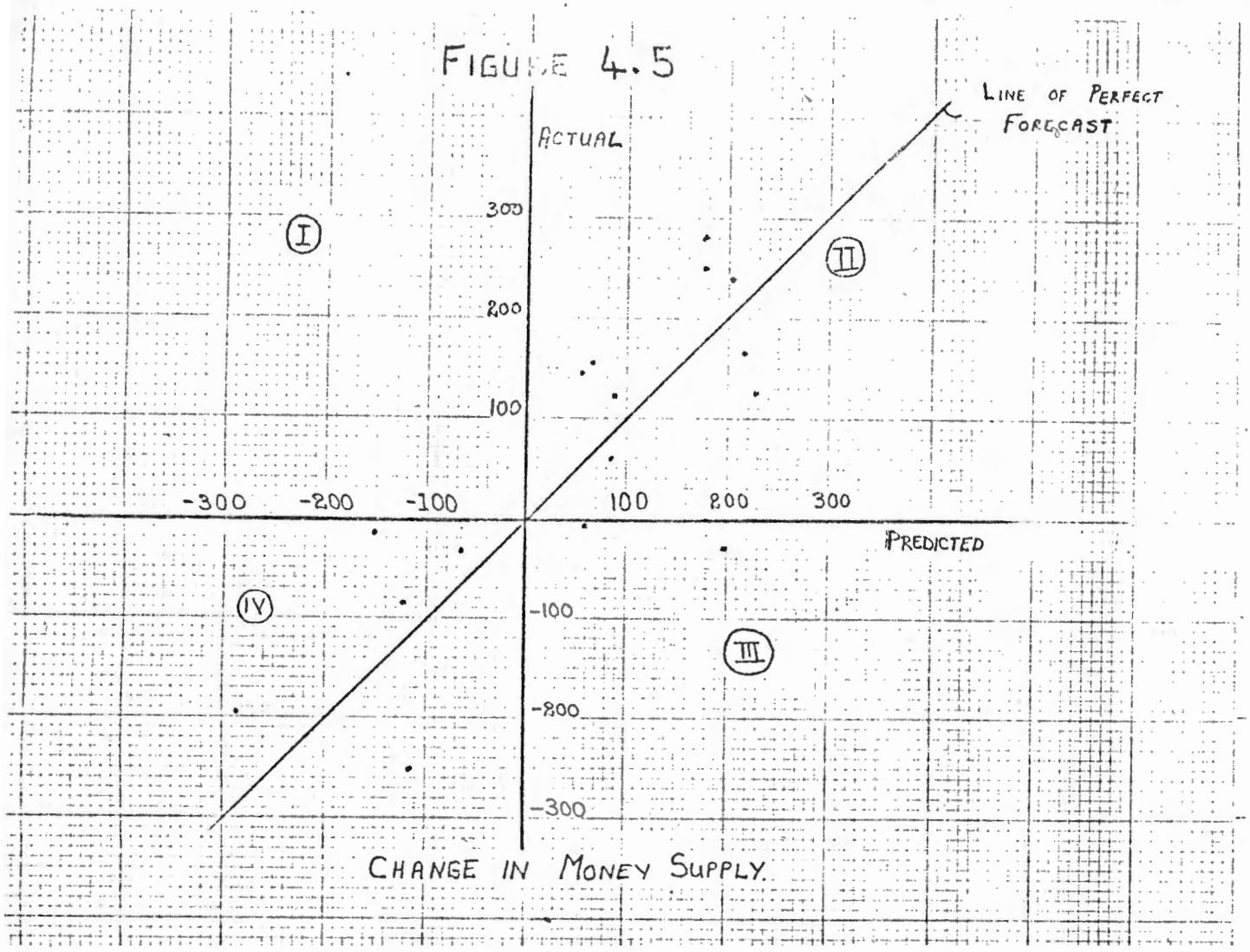
4.7.2 QUALITY OF FORECASTS

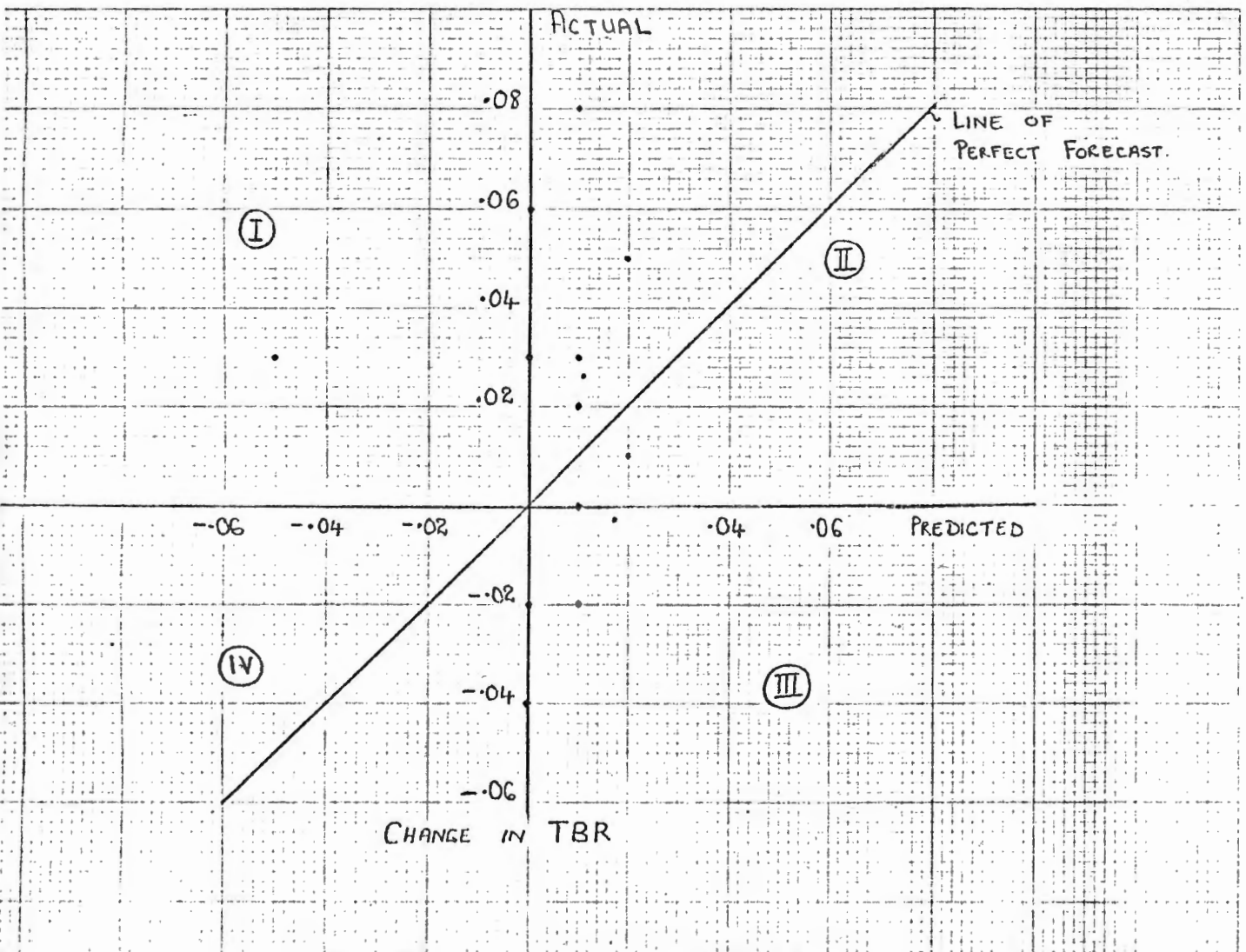
To assess the quality of the "best" forecasts we plot the actual change in a variable against the predicted change. This enables the analyser to determine if his model is able to predict the direction of change correctly. The univariate models for R and NDA are both random walk models and therefore the predicted change is zero at all times. Thus it is not possible to predict the variables R and NDA accurately. The plots for MS and TBR are given in Figure 4.5. The majority of plotted values for MS lie within the second or fourth quadrants which means that the changes in MS were predicted correctly. The direction of change for points lying in the first or third quadrants were predicted incorrectly. The forecasts produced for MS are satisfactory as most of them lie very close to the line of perfect forecast (it is when the actual change equals the predicted change). The quality of the forecasts produced for TBR are of a lower standard than those for MS as the plotted values are further away from the line of perfect forecast.

4.8 CONCLUSIONS

The different models discussed in this chapter provided good fits to the data, but in general the forecasts generated were poor. Necessarily the univariate models cannot incorporate information that might allow them to predict any kind of structural change in the economy as they extrapolate the behaviour of the economy during the estimation period.

FIGURE 4.5


LINE OF PERFECT FORECAST
ACTUAL
300
200
100
-300
-200
-100
100
200
300
PREDICTED
I
II
III
IV
-100
-200
-300
CHANGE IN MONEY SUPPLY


ACTUAL
·08
·06
·04
·02
LINE OF PERFECT FORECAST.
I
II
III
IV
-·06
-·04
-·02
·04
·06
PREDICTED
-·02
-·04
-·06
CHANGE IN TBR

Univariate forecasts should be used as a yardstick against which the forecasts from other methods can be compared. To adopt our forecasts for structural changes, we included explanatory variables in the estimated models. We find that various transfer function and regression models did not produce superior forecasts as the influence of the explanatory variables was usually out of line with the changes in the data. A re-estimation of the models using data which includes the forecast period may give non-significant coefficients so that these relationships do not hold any more.

Similar studies were done by Pierce (1977) and Cramer and Miller (1976). They also found a lack of relationships between several economic time series after prewhitening, which were causally related in economic theory. Cramer and Miller, however, had success with some causal relationships which gave an improvement on the univariate forecasts. More satisfactory results would be expected when systems, which do not change in structure are analysed so that the relationships extablished is expected to hold in the future.

The author plans to carry out further research on multiple time series modelling. Very little practical work has been done with time series models which have more than two independent variables.

A P P E N D I X A

THE DATA

The following monthly data series are taken from the South African Reserve Bank Quarterly Bulletins and their supplements unless otherwise stated. All monetary aggregates are expressed in Rand millions.

1. *Money* Base (MB) : Obtained by considering the Reserve Bank balance sheet Liabilities and excluding all Government and Provincial Administration deposits, Capital and Reserves, and "other Liabilities", and "Foreign Loans" (as they will be reflected in Reserves).

2. Foreign Reserves (R) : Total gold and foreign reserves of the Reserve Bank.

3. Net Domestic Assets (NDA) : Calculated as a residual MB-R.

4. Money Supply (MS) : Total Money and Near Money.

5. Treasury Bill Rate (TBR).

6. 90 day Negotiable Certificate of Deposits Rate (NCD).

7. The Euro-Dollar Rate (ED).
   (Source : International Monetary Fund, "International Financial Statistics.")

8. Electric Current Generated (E).
   (SARBQB 1970 = 100 diseasonalised.)

9. Government Deficit (gd) : We define "true government deficit" to equal the difference between government issues and the sum of government borrowing and government tax receipts.

In order to estimate the size of current government borrowing it was necessary to consider the change in the total holdings of government debt by

1. Permanent Building Societies,
2. Insurers (Long and Short Term)
3. Private Pension and Provident Funds
4. Banks (Commercial, Merchant and Hire Purchase), as well as by the Public Debt Commissioners and the Loan Levy Account.

For Permanent Building Societies we took (total) government borrowing equal to "Total Prescribed investments less Coin, Banknotes and Money at Call";

For the Insurers, (we aggregate the Long and Short term Insurers), total government borrowing is taken as equal to "Government Stock + Local Authority Stock + Public Corporation Stock + (approximated) Loans to Local Authorities" (where the last-named amount is assumed equal to ½ of "Other Loans").

Pension Funds holdings are taken equal to "Government Stock + Local Authority Stock + Public Corporation Stock + Loans to Local Authorities and to Public Corpor-

ations." For Unit trusts, take "(Total) Approved Securities."

In the Banking Sector, we take Commercial Bank holdings equal to ("Total other Prescribed Investments," less NCD's, plus "Other", plus "Treasury Bills, Bills and Advances to the Land Bank, Short-Term government stock and Short-Term debentures of the Land Bank") (up to October 1972, thereafter "Prescribed Investments excluding Liquid Assets").

Merchant Bank holdings are assumed equal to "Reserve Bank Balances" plus "S.T. government stock" plus "Other" liquid assets plus "Other" liquid assets plus (after October 1972) "Other" prescribed investments."

Hire-Purchase, Savings and General Banks holdings assumed equal to "Reserve Bank Balances" plus "S.T. Government stock" plus "S.T. Land Bank debentures" plus "Other government stock" plus "Local Authority and Public Corporation Stock."

Notes:

1. With respect to monetary banking institutions, after October 1972 "Liquid Assets" were not included with "Prescribed Investments."

2. We have excluded the Discount Houses in the above consideration in view of their purely intermediary role between financial institutions.

DATA NOVEMBER 1970 TO MARCH 1978

MB	R	NDA	MS	TBR	NCD	D	E	GD
850.00	745.00	105.00	3931.00	4.46	7.65	7.17	102.20	-6.00
856.00	681.00	175.00	3983.00	4.53	7.90	7.29	103.60	116.00
858.00	649.00	209.00	3874.00	4.64	8.60	5.93	104.60	5.00
842.00	641.00	201.00	3923.00	4.88	8.40	5.60	103.90	-77.00
859.00	637.00	222.00	3902.00	5.04	9.05	5.10	104.70	301.00
843.00	619.00	224.00	3963.00	5.51	8.00	5.95	104.70	142.00
838.00	611.00	227.00	3998.00	5.52	8.00	7.08	106.00	-87.00
843.00	570.00	273.00	4034.00	5.51	7.85	7.16	107.00	77.00
858.00	519.00	339.00	4088.00	5.48	7.90	6.46	110.30	141.00
838.00	486.00	352.00	4082.00	5.42	6.65	8.21	110.70	87.00
889.00	477.00	412.00	4151.00	5.54	8.20	8.46	110.30	109.00
876.00	460.00	416.00	4202.00	5.60	8.00	6.60	112.10	113.00
893.00	415.00	478.00	4219.00	5.77	8.35	6.28	111.10	113.00
818.00	464.00	354.00	4276.00	5.96	8.70	6.11	112.60	238.00
796.00	505.00	291.00	4118.00	5.95	8.45	5.37	111.80	129.00
765.00	520.00	265.00	4252.00	5.94	7.75	5.15	113.10	64.00
833.00	562.00	271.00	4185.00	5.94	8.30	5.28	112.70	192.00
835.00	616.00	219.00	4243.00	5.62	6.40	5.27	113.20	396.00
818.00	673.00	145.00	4276.00	5.55	6.50	4.88	115.20	327.00
629.00	732.00	97.00	4407.00	5.53	6.65	5.06	114.90	186.00
839.00	833.00	6.00	4440.00	5.32	6.20	5.58	115.90	184.00
823.00	839.00	-16.00	4509.00	4.50	5.00	5.49	113.80	202.00
870.00	858.00	12.00	4605.00	4.96	5.45	5.42	117.20	40.00
841.00	916.00	-75.00	4580.00	4.92	5.70	6.00	117.80	44.00
859.00	916.00	-57.00	4691.00	4.84	6.30	5.77	118.90	39.00
866.00	936.00	-70.00	4863.00	4.38	5.85	6.17	121.40	104.00
842.00	984.00	-142.00	4759.00	4.26	5.30	6.17	121.70	-9.00
848.00	998.00	-150.00	4893.00	4.28	6.00	7.45	123.00	41.00
867.00	1080.00	-213.00	5062.00	4.19	6.60	8.50	124.50	122.00

887.00	1138.00	-251.00	5272.00	3.34	4.50	8.16	125.70	278.00
888.00	1195.00	-307.00	5364.00	2.84	4.00	8.43	125.60	248.00
933.00	1196.00	-263.00	5446.00	2.64	4.50	8.81	123.70	171.00
933.00	1268.00	-335.00	5542.00	2.58	3.60	10.37	127.30	31.00
950.00	1176.00	-226.00	5558.00	2.48	5.30	11.46	127.90	-107.00
1005.00	1080.00	-75.00	5748.00	2.65	6.75	11.13	129.20	-105.00
980.00	1066.00	-86.00	5656.00	2.70	6.60	9.93	129.70	5.00
1070.00	924.00	146.00	5725.00	2.80	7.70	9.82	131.00	27.00
1058.00	796.00	262.00	5983.00	3.28	8.00	10.63	130.70	-39.00
1027.00	772.00	255.00	5881.00	4.68	9.00	9.36	132.60	-135.00
1061.00	835.00	226.00	6038.00	4.76	9.25	8.50	132.80	-203.00
1106.00	923.00	183.00	6161.00	4.82	9.00	9.23	134.70	596.00
1095.00	828.00	267.00	6191.00	4.82	10.00	10.53	136.30	319.00
1090.00	800.00	290.00	6391.00	4.82	9.75	11.67	136.00	-8.00
1122.00	777.00	345.00	6497.00	5.72	11.00	12.11	137.60	43.00
1119.00	784.00	335.00	6501.00	5.84	12.50	13.49	138.40	-85.00
1152.00	729.00	423.00	6438.00	6.36	14.25	13.56	140.90	-218.00
1218.00	737.00	481.00	6657.00	6.18	17.25	12.34	144.70	50.00
1179.00	754.00	425.00	6815.00	6.08	14.75	10.90	140.30	149.00
1242.00	817.00	425.00	7044.00	6.04	11.25	10.13	142.10	195.00
1241.00	729.00	512.00	7317.00	5.92	11.00	10.31	142.30	330.00
1222.00	740.00	482.00	7258.00	5.98	11.40	8.58	142.60	144.00
1255.00	747.00	508.00	7321.00	5.92	10.50	7.20	143.00	65.00
1292.00	751.00	541.00	7482.00	5.94	11.50	6.85	144.40	501.00
1269.00	714.00	555.00	7663.00	5.76	9.60	7.04	145.80	454.00
1294.00	744.00	550.00	7638.00	5.67	9.00	6.25	145.40	255.00
1297.00	720.00	577.00	7778.00	5.58	7.25	6.10	146.40	394.00
1314.00	763.00	551.00	7631.00	5.54	6.90	7.13	147.00	334.00
1358.00	794.00	564.00	7849.00	6.12	8.35	7.23	146.60	155.00
1466.00	933.00	533.00	8124.00	6.48	8.90	7.05	147.10	292.00

1469.00	972.00	497.00	8150.00	6.80	8.70	7.13	148.20	327.00
1510.00	972.00	538.00	8338.00	6.89	7.90	6.79	150.30	212.00
1499.00	940.00	559.00	8591.00	6.82	9.00	6.47	149.90	267.00
1561.00	960.00	601.00	8456.00	6.93	9.70	5.48	152.70	30.00
1569.00	859.00	710.00	8496.00	7.04	12.50	5.53	159.50	275.00
1482.00	1205.00	277.00	8740.00	7.17	13.50	5.60	153.00	-167.00
1500.00	1069.00	431.00	8954.00	7.21	11.50	5.41	148.10	363.00
1500.00	941.00	559.00	8966.00	7.19	12.00	5.96	152.50	437.00
1575.00	809.00	559.00	9213.00	7.31	12.50	6.22	150.20	726.00
1519.00	694.00	766.00	9116.00	7.49	11.20	5.79	154.10	81.00
1466.00	640.00	826.00	9076.00	7.76	11.93	5.67	150.40	393.00
1548.00	688.00	860.00	9256.00	7.76	10.69	5.57	153.60	-17.00
1523.00	703.00	820.00	9260.00	7.73	9.76	5.45	152.90	331.00
1532.00	740.00	792.00	9242.00	7.72	9.50	5.30	150.40	410.00
1500.00	734.00	766.00	9368.00	7.74	10.09	5.01	152.00	537.00
1522.00	719.00	803.00	9115.00	7.82	9.86	5.15	161.80	96.00
1525.00	645.00	880.00	9174.00	7.87	10.25	5.08	159.20	199.00
1484.00	649.00	835.00	9146.00	7.88	9.71	5.11	152.80	-125.00
1508.00	757.00	751.00	9427.00	7.86	9.45	5.13	146.20	566.00
1505.00	743.00	762.00	9548.00	7.82	9.05	5.77	151.90	722.00
1492.00	749.00	743.00	9799.00	7.85	8.94	5.78	151.90	491.00
1541.00	755.00	786.00	9677.00	7.88	8.73	5.77	155.00	453.00
1488.00	712.00	776.00	9613.00	7.88	8.66	6.30	154.10	284.00
1570.00	700.00	870.00	9853.00	7.86	9.00	6.56	154.10	235.00
1518.00	682.00	836.00	9703.00	7.92	8.90	7.13	166.40	554.00
1585.00	660.00	925.00	9848.00	7.92	8.70	7.08	164.00	.00
1664.00	636.00	958.00	10012.00	7.90	8.78	7.12	165.30	.00
1532.00	623.00	905.00	9726.00	7.93	9.06	7.31	165.10	.00
.00	631.00	899.00	9719.00	7.95	9.70	7.27	167.00	.00
.00	634.00	1026.00	.00	7.98	9.59	7.26	161.20	.00

A P P E N D I X B

COMPUTER PROGRAMS

The computer program COMET was used to solve the regression equations in Chapter 4. Comet is a UNIVAC Package that gives linear least-squares estimates of the regression equation coefficients.

The Box-Jenkins package written by David J. Pack of the Ohio State University was used for the analysis of time series in Chapter 4.

B I B L I O G R A P H Y

1. ANDERSON, O.D. (1977) : A Box-Jenkins Analysis of the Coloured Fox Data from Nain, Labrador. The Statistician Vol 26, No. 1

2. ANDERSON, O.D. (1975) : Time Series Analysis and Forecasting: The Box-Jenkins approach. Butterworths, London

3. ANDERSON, R.L. (1942) : Distribution of the serial correlation coefficient. Ann. Math. Stat. 13, 1-13.

4. ANDERSON, T.W. (1971) : The statistical Analysis of Time Series. Wiley, New York.

5. BARTLETT, M.S. (1946) : On the theoretical specification of sampling properties of autocorrelated time series. J. Roy. Stat. Soc. B, No. 8, 27-41.

6. BIRKHOFF, G.D. (1931) : Proof of the ergodic theorem. Proc. Nat. Acad. Sci., U.S.A., 17, 656.

7. BLOOMFIELD, P. (1972) : On the error of prediction of a time series. Biometrika, Vol. 59, No. 3.

8. BLOOMFIELD, P. (1976) : Fourier Analysis of Time Series: An Introduction. Wiley, New York.

9. BOX, G.E.P. and JENKINS, G.M. (1970) : Time Series Analysis, Forecasting and Control. Holden Day, San Francisco

10. BOX, G.E.P. and COX, D.R. (1964) : An analysis of transformations. J. Roy. Stat. Soc. B, 26, 211-243.

11. BOX, G.E.P. and JENKINS, G.M. (1973) : Some comments on a paper by Chatfield and Prothero and on a Review by Kendall. J. R. Statist. Soc. A, 136, Part 3, 337-352.

12. BOX, G.E.P., JENKINS, G.M. and MACGREGOR, J.F. (1974) : Some recent advances in forecasting and control II. Applied Statistics, Vol. 23, No. 2, 158.

13. BOX, G.E.P. and PIERCE, D.A. (1970) : Distribution of Residual Autocorrelations in ARIMA time series models. J.A.S.A., 65, 1509-1526.

14. CHATFIELD, C and PROTHERO, D.L. (1973) : Box-Jenkins Seasonal Forecasting : Problems in a case-study. J.R. Statist. Soc. A, 136, Part 3, 295-336.

15. CHATFIELD, C. (1977) : Some Recent Developments in Time Series Analysis. J.R. Statist. Soc. A, 140 Part 4, 492-510.

16. CHRIST, C.F. (1975) : Judging the performance of Econometric Models of the U.S. Economy. International Economic Review, Vol. 16 No. 1, 54-74.

17. CRAMER, R.H. and MILLER, R.B. (1976) : Dynamic Modelling of Multivariate Time Series for Use in Bank Analysis. Journal of Money, Credit and Banking, 8, 85-96.

18. DAVIES, N, PATE, M.B., FROST, M.G. (1974) : Maximum autocorrelations for moving average processes. Biometrika, 61, 199-200.

19. DHRYMES, P.J. (1970) : Econometrics. Harper and Row, New York.

20. FULLER, W.A. (1976) : Introduction to Statistical Time Series. Wiley, New York.

21. GILCHIRST, W. (1974) : Statistical Forecasting - The State of the Art. Omega, Vol. 2, No. 6, 733-750.

22. GRANGER, C.W.J. (1969) : Investigating Causal Relations by Econometric Models and Cross Spectral Methods. Econometrica, 37, 424-438.

23. GRANGER, C.W.J. and MORRIS, M.J. (1976) : Time Series Modelling and Interpretation. J.R. Statist. Soc. A, 139, Part 2, 246-257.

24. GRANGER, C.W.J. and NEWBOLD, P. (1973) : Some comments on the evaluation of economic forecasts. Appl. Econ. 5, 35-47.

25. GRANGER, C.W.J. and NEWBOLD, P. (1976) : Forecasting transformed series. J.R. Statist. Soc. B, Vol. 38, No. 2 189-203.

26. GRANGER, C.W.J. and NEWBOLD, P. (1977) : Forecasting Economic Time Series. Academic Press.

27. HANNAN, E.J. (1960) : Time Series Analysis. Methuen, London.

28. HANNAN, E.J. (1970) : Multiple Time Series, Wiley, New York.

29. HAUGH, L.D. (1976) : Checking the independence of two covariance stationary time series : A Univariate residual cross-correlation Approach. J.A.S.A. Vol. 71, No. 354, 378-385.

30. HAUGH, L.D. and BOX, G.E.P. (1977) : Identification of Dynamic Regression (Distributed Lag) Models connecting Two Time Series. J.A.S.A. Vol. 72, No. 357, 121-130.

31. HURWITZ, A.M. (1977) : An econometric analysis of South African Monetary Phenomena. Unpublished Ph.D. thesis, University of Cape Town.

32. JENKINS, G.M. and WATTS, D.G. (1968) : Spectral Analysis and its Applications. Holden Day, San Francisco

33. JENKINS, G.M. (1973) : Time Series Forecasting : Part I : Models for forecasting non-stationary time series. Part II : Model Building - Some Examples. Iscol and Lancord Conference on Forecasting.

34. JOHNSTON, J. (1972) : Econometric Methods. McGraw-Hill, New York.

35. KENDALL, M.G. (1973) : Time Series Analysis. Griffin, London.

36. KHINTCHIN, A.Y. (1932) : Zu Birkchoff's Losung des Ergodenproblems. Math. Ann. 107, 485.

37. KOUTSOYIANNIS, A. (1973) : Theory of Econometrics. MacMillan.

38. LEVENBACH, H. and REUTER, B.E. (1976) : Forecasting Trending Time Series with Relative Growth Rate Models. Technometrics, Vol. 18, No. 3, 261-272.

39. MAKRIDAKIS, S. and WHEELWRIGHT, S.C. (1978) : Interactive Forecasting. Holden Day.

40. NAYLOR, T.H., SEAKS, T.G., WICHERN, D.W. (1972) : Box-Jenkins method : An alternative to econometric models. International Statistical Review, Vol. 40, No. 2.

41. NELSON, C.R. (1973) : Applied Time Series Analysis for Managerial Forecasting. Holden Day, San Francisco.

42. NEWBOLD, P. (1973) : Bayesian Estimation of Box-Jenkins Transfer Function Models. J.R. Statist. Soc. B, Vol. 35, No. 2, 323-336.

43. NEWBOLD, P. (1974) : The exact likelihood function for a mixed autoregressive moving average process. Biometrika, 61, Part 3, 423-426.

44. NEWBOLD, P. and GRANGER, C.W.J. (1974) : Experience with Forecasting Univariate Time Series and the Combination of Forecasts. J.R. Statist. Soc. A, 137, Part 3, 131-165.

45. NEWBOLD, P. (1975) : The principles of the Box-Jenkins approach. Operations Research Quarterly, Vol. 26. No. 2(ii).

46. PIERCE, D.A. (1972) : Residual Correlations and Diagnostic Checking in Dynamic Disturbance Time Series Models. J.A.S.A., Vol. 67, No. 339, 636-640.

47. PIERCE, D.A. (1975) : Forecasting in Dynamic Models with Stochastic Regressors. Journal of Econometrics, 3, 349-374.

48. PIERCE, D.A. (1975) : On Trend and Autocorrelation. Communication in Statistics, 4(2), 163-175.

49. PIERCE, D.A. (1977) : Relationships - and the Lack thereof - Between Economic Time Series, with Special Reference to Money and Interest Rates. J.A.S.A., Vol. 72, No. 357, 11-26.

50. PIERCE, D.A. and HAUGH, L.D. (1977) : Causality in Temporal Systems. Journal of Econometrics, 5, 265-293.

51. PROTHERO, D.L. and WALLIS, K.F. (1976) : Modelling Macroeconomic Time Series. J.R. Statist. Soc. A, 139, Part 4, 468-501.

52. QUENOUILLE, M.H. (1949) : Approximate tests of correlation in time series. J. Roy. Stat. Soc. B., 11, 68-84.

53. REID, D.J. (1973) : A Review of Short term Projection Techniques.

54. SIMS, C.A. (1972) : Money, Income and Causality. American Economic Review, 62, 540-552.

55. WICHERN, D.W. (1973) : The behaviour of the sample autocorrelation function for an integrated moving average process. Biometrika, 60, Part 2, 235-239.

56. YULE, G.U. (1927) : On a method for investigating periodicities in disturbed series with special reference to Wolfer's sunspot numbers. Phil. Trans. Roy. Soc. London, A 226, 267-298.

57. ZELLNER, A. and PALM, F. (1974) : Time Series Analysis and Simultaneous Equation Econometric Models. Journal of Econometrics, Vol. 2, 17-54.